UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1996

SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 2A

ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1996

20010319 008

AQM01-06-1187

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0727

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	
	December, 1996	

4. TITLE AND SUBTITLE
1996 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 2A, Armstrong Laboratory

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Armstrong Laboratory.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 2 is bound in two parts, 2A and 2B. Volume 2A contains #1-22. Volume 2B contains reports #23-43. The Table of Contents for Volume 2 is included in both parts.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A, 5B & 5C	Wright Laboratory
6	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center, and Wright Patterson Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

SFRP FINAL REPORT TABLE OF CONTENTS i-xviii

<u>APPENDICIES:</u>

SFRP FINAL REPORTS

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

iv

SRP Final Report Table of Contents

v

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After

completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	~117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/Mis annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant

4

increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.

GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)

HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	$690	$718	$740	$740	$740	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of $780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of $444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A -- PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

Lab	Evals Recv'd	How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
		0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
    a. Time available for evaluation of applications:
    b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5   a. Benefits of associate's work to laboratory:
    b. Benefits of associate's work to Air Force:
6.   a. Enhancement of research qualifications for LFP and staff:
    b. Enhancement of research qualifications for SFRP associate:
    c. Enhancement of research qualifications for GSRP associate:
7.   a. Enhancement of knowledge for LFP and staff:
    b. Enhancement of knowledge for SFRP associate:
    c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.   a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13.   a. Quality of RDL's communications with you:
       b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	AL	FJSRL	PL	RL	WHMC	WL
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.   a. Would you apply again?
      b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a.   At Home:
    b.   With Friend:
    c.   On Local Economy:
    d.   Base Quarters:
17. Value of orientation visit:
    a.   Essential:
    b.   Convenient:
    c.   Not Worth Cost:
    d.   Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

# res	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year?   Yes (92 %)
14. Will you pursue future studies related to this research?   Yes (68 %)
15. Was Tour length satisfactory?   Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

REDUCTION KINETICS IN A BATCH METALLIC IRON/WATER SYSTEM: EFFECT OF IRON/WATER EXPOSURE

Richelle M. Allen-King, Assistant Professor
Department of Geology

Washington State University
Pullman, WA 99164-2812

REDUCTION KINETICS IN A BATCH METALLIC IRON/WATER SYSTEM: EFFECT OF IRON/WATER EXPOSURE

Richelle M. Allen-King
Assistant Professor
Department of Geology
Washington State University
Pullman, WA 99164-2812

ABSTRACT

The reductive transformation rates of acetylene and carbon tetrachloride were observed in batch Fe^0/water systems. The pseudo-first order rate constants determined for both compounds were lower for "aged" (iron and water exposed prior to addition of the target analyte) relative to "fresh" (iron, water and the analyte mixed simultaneously) iron. The initial pseudo-first order rate constants were concentration-dependent, indicating nonlinear kinetics, in fresh iron Fe^0/water systems, while the pseudo-first order rate constants were independent of concentration in aged Fe^0/water systems. The rate of transformation slowed with time in the fresh Fe^0/water systems. Duration of iron exposure to water may be one contributor to the observed scatter in rate coefficients for individual compounds reported in the literature for Fe^0/water systems. An explanation consistent with the observations is that a low number of highly reactive "sites" present in the fresh iron are either blocked by precipitates or reaction products or eliminated by corrosion as the iron ages. Although initial reaction rates are classically used to determine reaction kinetics, initial rates in fresh Fe^0/water systems are not likely to hold practical relevance to applications proposed for groundwater and wastewater treatment. In treatment systems, the Fe^0 is continuously exposed to water.

REDUCTION KINETICS IN A BATCH METALLIC IRON/WATER SYSTEM: EFFECT OF IRON/WATER EXPOSURE

Richelle M. Allen-King

INTRODUCTION

Chlorinated solvents in groundwater are of public concern because they can have negative impacts on human health and the environment. Drinking water standards are typically in the low part-per-billion range. For example, the drinking water standards (maximum contaminant level or MCL) for trichloroethene (TCE), perchloroethene (PCE), and carbon tetrachloride (CTET) are 5 μg/L [US EPA, 1987; US EPA, 1991]. Chlorinated solvents are the most frequently identified organic contaminants in groundwater for several reasons: 1) these highly chlorinated compounds tend to be persistent in aerobic groundwater; 2) during the past 40-50 years, these compounds have been commonly used industrial solvents and they have been discharged to the environment by disposal, leakage and spillage, often as a dense non-aqueous phase liquid (DNAPL); 3) DNAPLs in the subsurface, distributed as both residual and pools, can provide exceptionally long-term sources of groundwater contamination that are challenging to delineate and remediate [Mackay and Cherry, 1989; Westrick, 1990; NRC, 1993; NRC, 1994; MacDonald and Kavanaugh, 1994; Vogel et al., 1987; Pankow et al., 1996].

The US Air Force (USAF) has identified at least 1100 sites contaminated with chlorinated solvents [Burris, pers. comm.]. Therefore, treatment methods for chlorinated-solvent contaminated sites

are of particular interest to the USAF. The conventional approach to groundwater remediation has been to pump contaminated groundwater to the surface for treatment [US EPA, 1990]. Because of the ineffectiveness of "pump-and-treat" as a remediation method, focus has shifted to alternative *in situ* remediation methods [NRC, 1993; NRC, 1994].

Zero-valent or metallic iron can bring about the reduction of dissolved chlorinated solvents such as TCE, PCE and CTET [Matheson and Tratnyek, 1994; Gillham and O'Hannesin, 1994 and references reviewed therein]. As a result, metallic iron has promise for use in both *in situ* and *ex situ* treatment of water contaminated with chlorinated solvents. The use of metallic iron in permeable *in situ* treatment barriers is being tested as an alternative remediation and containment method for chlorinated solvents in groundwater [Gillham, 1995; Wilson, 1995; among others]. Within this concept, contaminated groundwater flows naturally through a portion of the aquifer modified to contain iron, and the contaminants are remediated by reaction with the iron. Because transformation rates are rapid relative to typical groundwater flow velocities [Gillham and O'Hannesin, 1994; Sivavec and Horney, 1995], the *in situ* method has potential to treat groundwater with a passive system that may prove cost-effective relative to the conventional method [Wilson, 1995]. *Ex situ* treatment using an above-ground reactor has been demonstrated at a superfund site as part of the US Environmental Protection Agency SITE program. In either the *in situ* or *ex situ* applications, optimization and assurance of effective treatment system design relies on an understanding of the factors controlling the transformation reactions.

In a recent compilation of literature, Johnson et al. [1996] report that normalized rate coefficients

(k_{SA}, normalized for iron surface area to solution ratio) vary by about 1 order of magnitude for a particular compound in metallic iron/water (Fe^0/H_2O) systems. A broad range of variables are suggested as potentially important by these researchers [Johnson et al., 1996], including, concentration, treatment and handling of the iron samples, interpretation of rate constants, and a number of solution effects - such as pH, alkalinity, and dissolved oxygen, among others.

The effects of variables such as concentration and duration of iron/water exposure (aging) prior to contaminant contact have been reported in few studies. In the manuscript described above, Johnson et al. [1996] provide an example of a concentration-dependent initial transformation rate. The experiments described focus on carbon tetrachloride (CTET) present at high solute concentration:Fe^0 surface area (SA) ratios (Table 1). In contrast, Orth and Gillham [1996] have found that the TCE transformation can be modelled by pseudo-first order kinetics over a much lower range of solute concentration:Fe^0 SA ratios (Table 1). Much of the evidence concerning the effects of short-term aging in Fe^0/water systems on reaction rates is anecdotal [e.g. Matheson and Tratnyek, 1994; Sivavec and Horney, 1995]. These two variables are important because they provide insights into the fundamental chemistry of the reaction, as well as practical information needed for treatment system design.

Table 1. *Solute concentration: Fe⁰ surface area (Fe⁰ SA) estimated for experiments where rates were determined for a range of concentrations.*

Compound	Experiment Method	Estimated Solute Concentration: $Fe^0 SA^\dagger$ [μmol/m^2]	Fe^0 SA:H$_2$O [m^2/mL]	Experimental Data
CTET	Batch	20 - 780	0.001	Johnson et al., 1996
TCE	Column	0.036 - 2.2	0.21	Orth and Gillham, 1996

† Estimated by the author from information provided in the referenced manuscripts.
CTET is carbon tetrachloride; TCE is trichloroethene.

In this study, we document the effect of both aging and varying concentration on the kinetics of two relatively "fast" reacting compounds - acetylene and carbon tetrachloride (CTET) in batch Fe⁰/water systems. Acetylene is a key intermediate observed in the degradation of di-, tri-, and tetrachloroethene transformation in Fe⁰/water systems [Sivavec and Horney, 1995; Allen-King et al., in press; Roberts et al., 1996; Campbell and Burris, 1996]. The results of this study will improve our overall understanding of chlorinated solvent reactions in Fe⁰/H$_2$O systems and provide important practical information for treatment system design.

MATERIALS AND METHODS

Iron and pyrite

Electrolytic iron (Fisher Scientific) was used for this study. All iron was pretreated in 400 g batches by washing with 1 L of 1 N HCl for one-half hour then rinsing with 10 L of argon-sparged deionized water. After rinsing, the iron was dried overnight under a nitrogen atmosphere at 90°C. Once dried, the iron was stored in argon sparged glass jars to inhibit oxidation. Pyrite, obtained from Wards Natural Science Establishment, was used to pH buffer the batch system. Pretreatment of the pyrite consisted of powdering with a mortar and pestle and storing in argon sparged glass jars.

Batch system

The batch systems were prepared following the general method outlined by Roberts et al. (1996) and Campbell and Burris (1996). The reactive batch systems contained 5.00 g of pretreated iron and 0.10 g of powdered pyrite in argon-purged 150 ml nominal (~160 ml observed) Wheaton serum vials with 100 mL milliQ water and sufficient glass beads to result in a final headspace volume equal to 55 mL. Pyrite was added for consistency with prior experiments and to maintain the pH [Burris et al., 1995; Lipczynska-Kochany et al., 1994; Roberts et al., 1996]. For the "fresh" reactive systems (iron and water not mixed prior to exposure to compound of interest), the iron and pyrite were sealed under vacuum in 2 mL borosilicate glass ampules. Following

equilibration at 20 °C, the ampule was broken to initiate the reaction. For the lower acetylene concentrations (≤ 2 μmol) and all CTET experiments, the analyte (acetylene or CTET) and pentane internal standard (for the acetylene experiments) [Campbell and Burris, 1996], were added to the system as gasses. For the high concentration acetylene experiments, the de-aired milliQ water added to the system was pre-equilibrated with acetylene gas (0.999%) prior to addition. In either case, water was dispensed to the vials in an anaerobic glove box (5% H_2, balance N_2 atmosphere) to prevent oxygen from entering the headspace, and the vials were sealed with teflon-lined rubber septa (Supelco). An iron- and pyrite-free control vial was prepared and analyzed with each reactive vial to check for system losses attributable to mechanisms other than reduction. After preparation, vials were rotated at 6 rpm at 20 °C and concentrations of the reactants and products in the headspace were monitored over time.

Analytical

Concentrations were determined by collecting a 100 or 200 μL subsample of the headspace for acetylene, or a 10 μL subsample for CTET, and analyzing directly by gas chromatography. Acetylene and C_1-C_6 transformation products were analyzed on a Hewlett-Packard Model 5890 gas chromatograph (HP5890 GC) equipped with flame ionization detector, and a 30 m megabore GSQ-PLOT (J&W Scientific) capillary column. CTET and transformation products, chloroform and dichloromethane, were analyzed using a HP5890 GC, He carrier gas, 0.53 mm ID x 15 m DB-1 (J&W Scientific) column, equipped with electron capture detector.

Calibration was by the internal standard method for acetylene and by the external standard method

for CTET. Calibrations were conducted on a total mass basis using standards containing the parent and expected transformation products at three concentrations; standards contained 100 mL water and sufficient glass beads to equal the volume of the iron, broken ampule and pyrite in the reactive vials resulting in a headspace volume of 55 mL. The C_3-C_6 hydrocarbon products were quantified from the response factors for propane, butane, pentane, and hexane [Campbell and Burris, 1996]. Product sorption was neglected because of the very low carbon content of the electrolytic iron (Burris et al., in prep.).

RESULTS AND DISCUSSION

The initial reaction rate for acetylene in a fresh Fe^0/H_2O system was much more rapid than in an aged (iron exposed to water prior to addition of the acetylene) Fe^0/water system (Fig. 1). The initial pseudo-first order rate constants diminished by more than an order of magnitude with pre-exposure between iron and water (aging) for as little as one day.
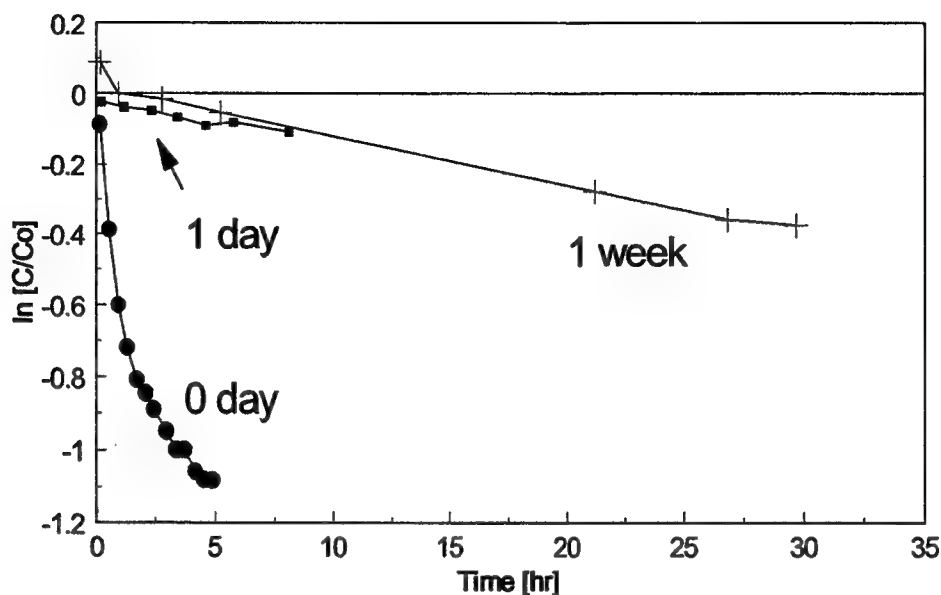
Fig. 1. *Acetylene (0.1 μmol) transformation in batch systems with fresh (no pre-exposure between iron and water prior to acetylene addition) and aged (Fe⁰ exposed to water prior to acetylene addition for the time period noted) iron.*

The initial pseudo-first-order rate constant for acetylene in fresh Fe^0/H_2O systems was concentration dependent. Faster disappearance was observed at lower concentrations (Fig. 2, Table 1). For all of the concentrations shown, the transformation rate declined with time, apparently trending asymptotically towards a slower rate, comparable to the rates observed in the aged Fe^0/H_2O systems. The change in rate with time was most pronounced for lower acetylene concentrations tested.
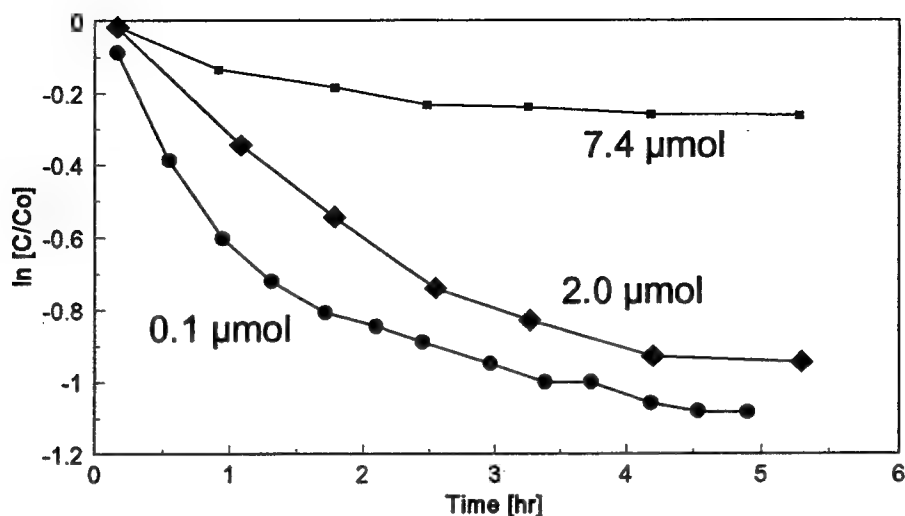
Fig. 2. *Acetylene transformation in batch systems with fresh iron (no pre-exposure between iron and water prior to acetylene addition).*

The trends in behavior of CTET were similar to those observed for acetylene: pseudo-first order reaction rates were more rapid in fresh Fe^0/H_2O systems than in aged systems (Table 2). The transformation rate declined with time in the fresh Fe^0/H_2O system; and the observed rate constant seemed to trend asymptotically towards the rate constant observed in the aged Fe^0/H_2O systems.

In Fe^0/H_2O systems aged for 3 to 7 d, the pseudo-first-order rate constants observed were apparently not concentration dependent for either CTET or acetylene (Table 2). It is also interesting to note that pseudo-first-order rate constants observed for CTET and acetylene in the fresh Fe^0/H_2O systems appear to be consistent for a comparable concentration.

Table 2. *Initial pseudo-first-order rate constants [hr^{-1}] for acetylene or CTET transformation in fresh and aged Fe^0/water systems.*

Compound	Initial Mass [μmol]	pseudo-first-order rate constants [hr^{-1}] for different Fe^0/water aging period	
		fresh	3 to 7 days
Acetylene	0.1	0.55	0.015
Acetylene	2.0	0.30	0.013
CTET	0.1	0.60	0.33
CTET	1.0	0.45	0.30

The results presented above suggest that the initial reaction rate observed may be controlled by a limited number of very fast reacting sites which are eliminated within a few hours in the experiment systems. A limited number of such sites is consistent with the apparent concentration effect on the initial rates in the fresh Fe^0/H$_2$O systems in this study. The fact that consistent concentration dependent behavior was observed for both compounds in the fresh Fe^0/H$_2$O systems may be an indication that the rate-limiting step for the initial rate is transport to the fast reacting-sites. The fast sites could be eliminated by any of several physical mechanisms, including: blockage by precipitates or irreversible sorption of reaction products, or corrosion. Other researchers have observed diminishing reaction rates with time in batch systems [Gillham and O'Hannesin, 1994; Johnson et al., 1996; Allen-King et al., in press].Although initial reaction rates are classically used to determine reaction kinetics, initial rates in fresh Fe^0/water systems are not likely to hold practical relevance to applications proposed for groundwater and wastewater treatment. In treatment systems, the iron is continuously exposed to water.

1-12

Johnson et al. [1996] have reported concentration-dependent CTET transformation kinetics in batch Fe^0/water systems with relatively high CTET concentration:Fe^0 surface area ratios. The CTET concentration:Fe^0 surface areas for the current are estimated to be more than 100 times lower than the ratios reported by Johnson et al. [1996].

The C mass balances in all the fresh Fe^0/water acetylene experiments are poor. Fig. 3 provides an example of typically observed behavior; approximately 60 to 70% of the C was recovered in these systems. Poor mass balances for acetylene in fresh Fe^0/water systems have also been observed by Burris (pers. comm.). The principal transformation products observed were ethene, ethane and C_4 hydrocarbons. Much lower concentrations of methane and C_3 compounds were observed. Most of the mass balance error was generated during the early part of the experiment, corresponding to the early period when the disappearance rate was most rapid. These results seem to suggest a connection between the poor mass balance and initial rapid disappearance, such as an irreversible sorption or polymerization step. The C balances were nearer to 100% in the experiments in aged Fe^0/water systems. However, the proportion of acetylene transformed in these experiments was also less than in the fresh Fe^0/water systems.
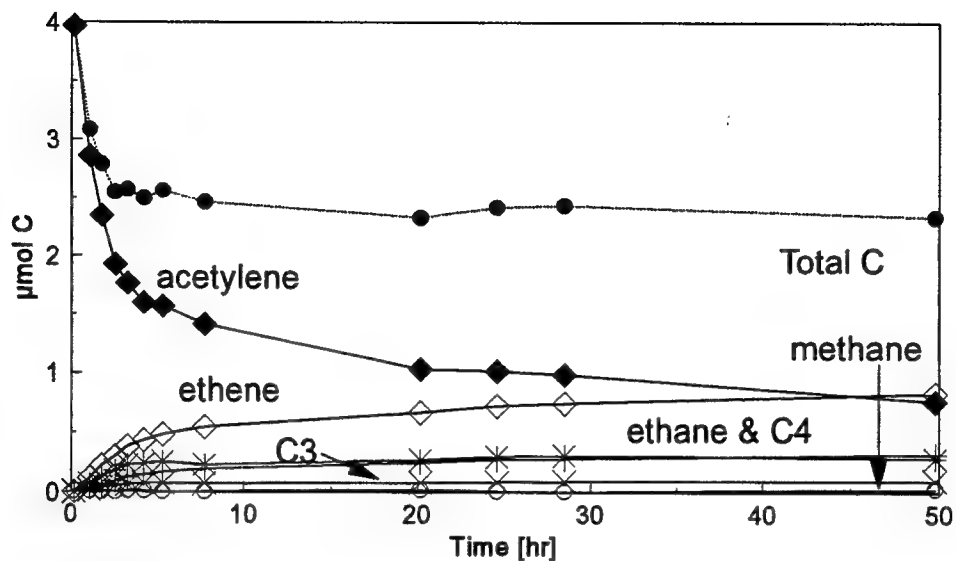
Fig. 3. *Acetylene (2 μmol) and degradation products in a batch system with fresh iron (no pre-exposure between iron and water prior to acetylene addition).*

REFERENCES

Allen-King, R. M., R. M. Halket, and D. R. Burris. In press. Reductive transformation and sorption of cis- and trans-1,2-dichloroethene in a metallic iron/water system. *Environ. Toxicol. Chem.*

Burris, D. R., T. J. Campbell, and V. S. Manoranjan. 1995. Sorption of trichloroethylene and tetrachloroethylene in a batch reactive metallic iron-water system. *Environ. Sci. Technol.* 29:2850-

2855.

Burris, D. R., V. S. Manoranjan, and T. J. Campbell. In prep. Tetrachloroethylene and trichloroethylene reduction by metallic iron: Sorption and mass transfer limitations. *J. Environ. Engin.*

Campbell, T. J. and D. R. Burris. 1996. Analysis of chlorinated ethene reduction products in vapor/water phase systems by dual-column, single-detector gas chromatography. *Intern. J. Environ. Anal. Chem.* 63:119-126.

Gillham, R. W., and S. F. O'Hannesin. 1994. Enhanced degradation of halogenated aliphatics by zero-valent iron. *Ground Water* 32:958-67.

Johnson, T. J., M. M. Scherer, and P. G. Tratnyek. 1996. Kinetics of halogenated organic compound degradation by iron metal. *Environ. Sci. Technol.* 30:2634-2640.

Lipczynska-Kochany, E., S. Harms, R. Milburn, G. Sprah, and N. Nadarajah. 1994. Degradation of carbon tetrachloride in the presence of iron and sulphur containing compounds. *Chemosphere,* 29:1477-1489.

MacDonald, J. A., and M. C. Kavanaugh. 1994. Restoring contaminated groundwater: An achievable goal? *Environ. Sci. Technol.* 28:362-68.

Mackay, D. M., and J. A. Cherry. 1989. Groundwater contamination: Pump-and-treat remediation. *Environ. Sci. Technol.* 23, 6:630-36.

Matheson, L. J., and P. G. Tratnyek. 1994. Reductive dechlorination of chlorinated methanes by iron metal. *Environ. Sci. Technol.* 28:2045-53.

NRC (National Research Council). 1993. *In Situ Bioremediation: When Does it Work?* Washington, D.C.: National Academy Press.

NRC (National Research Council). 1994. *Alternatives for Groundwater Cleanup.* Washington, D.C.: National Academy Press.

Orth, W. S. and R. W. Gillham. 1996. Dechlorination of trichloroethene in aqueous solution using Fe^0. *Environ. Sci. Technol.*, 30:66-71.

Pankow, J. F., S. Feenstra, J. A. Cherry, and M. C. Ryan. 1996. Dense chlorinated solvents in groundwater: Background and history of the problem. **In** *Dense Chlorinated Solvents and Other DNAPLs in Groundwater.* J. Pankow, and J. A. Cherry (eds.), 1-52. Portland, OR: Waterloo Press.

Roberts, A. L., L. A. Totten, W. A. Arnold, D. R. Burris, and T. J. Campbell. 1996. Reductive elimination of chlorinated ethylenes by zero-valent metals. *Environ. Sci. Technol.* 30:2654-2659.

Sivavec, T. M., and D. P. Horney. 1995. Reductive dechlorination of chlorinated ethenes by iron metal. *209th ACS National Meeting, Anaheim, CA* 35, 1 (April 2-7):695-98.

US EPA (Environmental Protection Agency). 1987. *Federal Register,* July 8, 1987. Washington, D.C.: U. S. Government Printing Office.

US EPA (Environmental Protection Agency). 1990. A Guide to Pump and Treat Groundwater Remediation Technology. EPA/540/2-90/018.

US EPA (Environmental Protection Agency). 1991. *Federal Register,* January 30, 1990. Washington, D.C.: U. S. Government Printing Office.

Vogel, T. M., C. S. Criddle, and P. L. McCarty. 1987. Transformations of halogenated aliphatic compounds. *Environ. Sci. Technol.* 21:722-36.

Westrick, J. J. 1990. National surveys of volatile organic compounds in ground and surface waters. In *Significance and Treatment of Volatile Organic Compounds in Water Supplies* N. M. Ram (ed.), 103-25. Chelsea, MI: Lewis.

Wilson, E. K. 1995. Zero-valent metals provide possible solution to groundwater problems. *Chemical & Engineering News,* July 3, 19-22.

INVESTIGATION OF THE ELECTROCHEMILUMINESCENT PROPERTIES
OF SEVERAL NATURAL AND SYNTHETIC COMPOUNDS

Anthony Andrews
Assistant Professor
Centre for Intelligent Chemical Instrumentation

Clippinger Laboratories
Ohio University
Athens, OH, 45701-2979

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

August 1996

INVESTIGATION OF THE ELECTROCHEMILUMINESCENT PROPERTIES OF SEVERAL NATURAL AND SYNTHETIC COMPOUNDS

Anthony Andrews
Assistant Professor
Centre for Intelligent Chemical Instrumentation
Ohio University

Abstract

The fluorescence properties of a number of compounds of interest to the US Air Force were studied. Some of these compounds have been shown to give rise to electrochemiluminescence (ECL) under certain conditions. The fluorescence change was monitored over a period of time for these electrochemiluminescent compounds.

An electrochemiluminescent analysis system with spectral capabilities was constructed in-house from a photon counting spectrofluorometer and an electrochemical workstation (potentiostat). The purpose of this instrument was to study the ECL properties of several naturally occurring compounds such as extracted body fluids from tunicates, synthetic analogue compounds such as 2,3,4-tunichrome and 3,4,5-tunichrome and diaminotoluenes in the presence and absence of metal ions.

By combining these two instruments the intention was to probe the electrochemical reactions and obtain ECL spectra for a number of compounds. These spectra could then be compared with fluorescent spectra to help identify emitting species. Synthetic ECL compounds were used as models of naturally occurring biological ECL systems to further elucidate the ECL reaction mechanisms.

INVESTIGATION OF THE ELECTROCHEMILUMINESCENT PROPERTIES OF SEVERAL NATURAL AND SYNTHETIC COMPOUNDS

Anthony Andrews

Introduction

ECL is the generation of light during a chemical reaction with at least one of the reagents being generated *in situ* at an electrode. This has the advantage of allowing the reaction to be controlled and managed by adjustments to the applied potential. The occurrence, mechanisms and analytical applications of ECL are well covered in a recent review (1). Probably the most common electrochemiluminescent reagent is tris(2,2'-bipyridinyl)ruthenium(II) [$Ru(bpy)_3^{2+}$] which has been used to determine a number of different compounds including aliphatic amines (2), drugs (3), amino acids (4) and oxalate in urine (5).

Perhaps the greatest potential for ECL lies in the area of sensors. Electrochemical sensors have the potential for analytical measurement to move from the laboratory to the field, and to perform these measurements rapidly, inexpensively and reliably.

By utilising the additional selectivity of ECL it is feasible that a major disadvantage of field electrochemical sensors could be overcome, that is, their lack of selectivity. This is because the addition of a chemiluminescence step adds extra selectivity to the sensor in addition to the discrimination provided by the electrochemical stage.

The use of naturally occurring compounds, or their synthetic analogues, offers the highest selectivity. These sensors are known as biosensors and utilise the extreme selectivity frequently found inherently in biological molecules such as antibodies, enzymes and proteinaceous receptors.

In order for biosensors to become a reality, more fundamental electrochemiluminescent reactions must be discovered and characterised. One such group of compounds that has the potential for use in ECL biosensors are found in tunicates. Tunicates are sessile filter feeding marine animals commonly called "sea squirts" which produce a family of blood cell pigments referred to as "tunichromes". Tunichromes bind Fe^{3+} and many other metals, which can be determined by their ability to concentrate certain metals far above ambient seawater concentrations (6).

Although initial research results concerning the ECL of tunicates and synthetic analogues of compounds found in their blood, the tunichromes, have been published (7), there still remains the need for a detailed analysis of the ECL reactions occurring and the determination of the emitting species. Elucidation of the emitting species in this complex reaction may open the door for more efficient synthetic compounds to be produced. In addition with knowledge of the emitting species the first step towards construction of ECL biosensor will have been taken as the selectivity, specificity and sensitivity of the reaction can then be quantitated.

Experimental

Chemicals

A stock electrochemical test solution of potassium ferricyanide (1×10^{-2} M, ACS grade, Fisher) in potassium nitrate (1 M, Fisher) was prepared and cyclic voltammagrams from this solution were used to test the electrochemical system (8,9).

Synthetic tunichrome analogues, N,N',N''-tris(2,3,4-trihydroxybenzoyl)-1,5,10-triazadecane (I); hereafter referred to as "2,3,4-T" and N, N',N''-tris(3,4,5-trihydroxybenzoyl)-1,5,10-triazadecane (II); hereafter referred to as "3,4,5-T" (10) were the generous gifts of Dr. Carl Carrano (Dept. of Chemistry, Southwest Texas State University, San Marcos, TX.). Solutions of 2,4-diaminotoluene (2,4-DAT), 3,4-diaminotoluene (3,4-DAT) and 3,4-dinitrotoluene (3,4-DNT) (all Aldrich) were prepared at the 1 mg/mL level in methanol (HPLC grade, Fisher). All water used was distilled and deionized prior to use. Phosphate buffered saline (PBS) was prepared by dissolving sodium dihydrogen phosphate monohydrate (0.1284 g), disodium hydrogen phosphate (0.6002 g) and sodium chloride (4.3994 g) in water (500.0 mL). A solution of $Ru(bpy)_3^{2+}$ (1×10^{-3} M, Aldrich) in PBS was prepared and used as an ECL test compound. Tripropylamine (TPA) solutions used were assay buffer (Igen Corp, MD) used as received.

I
R = OH
R_1=OH
R_2=OH
R_3=H

II
R = H
R_1=OH
R_2=OH
R_3=OH
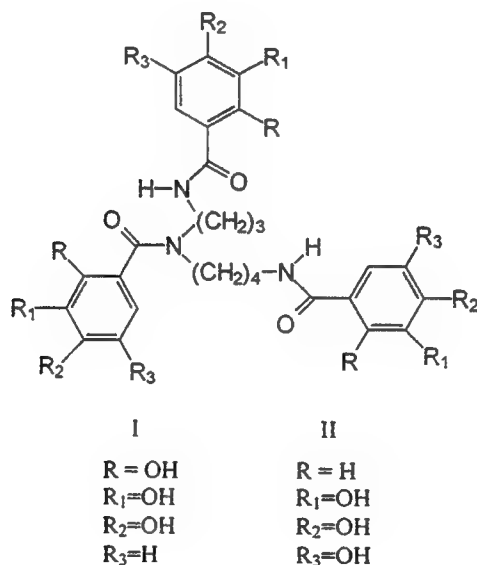
Samples used previously for ECL studies are detailed in Table 1. Each sample was given a number as it was logged in to aid in identification and data acquisition on the computer data files generated.

Table 1 List of samples used previously in ECL studies and used in this research.

#	Date	Sample	Volume /mL	Colour
1	No date	3,4-DAT in MeOH	<5	rust-brown
2	3/18/96	TL 2,3,4-T	15-20	pale yellow
3	5/22/96	Prot. K. M. Occ. plasma	<5	pale yellow
4	7/3/96	1,2-DAQ	< 5	dark purple
5	3/18/96	3,4,5-T (0.2 mg/mL in MeOH:TPA 1:15	15-20	clear
6	no date	2,3-DAT in MeOH 2 mg/mL	5-10	rust-brown
7	10/7/95	0.5 mg/mL hematin in KOH (1 M)	<5	dark green
8	4/11/96	bipyridine 1 mg/mL in MeOH	15	clear
9	5/22/96	M. Occ. blood + 0.5% protease K	<5	yellow
10	7/12/96	3,4-DAT in MeOH 1 mg/mL	5	gold
11	3/18/96	TL 3,4,5-T + Tl	15	clear
12	3/20/96	Ru(bpy)$_3$	<5	red-orange
13	no date	0.3 M TPA assay buffer	10-15	clear
14	7/10/96	3,4-DAT 1 mg/mL in MeOH (pooled with 7/3/96)	<5	gold
15	5/14/96	M. Occ. blood	45	opaque
16	7/3/96	3,4,5-T 1 mg/mL in MeOH	<5	clear
17	3/18/96	2,3,4-T in mL MeOH? + 100 ppm K	15	clear
18	7/17/96	M. Occ. blood cell lysate	<1	clear
19	7/17/96	M. Occ. blood cell/Protease K	<1	lt. brown
20	7/17/96	M. Occ. blood serum	30	lt. brown
21	3/96	vanadium + 2,3,4-T	powder	brown
22	7/21/96	2,4-DAT in methanol (1 mg/mL)	50	clear
23	7/21/96	3,4-DAT in methanol (1 mg/mL)	50	light brown
24	7/21/96	3,4-DNT in methanol (1 mg/mL)	50	faint green
25	7/21/96	3,4,5-T in methanol (1 mg/mL)	50	clear

Equipment

The electrochemical experiments were performed using a Model 660 electrochemical workstation (CH Instruments, Memphis TN). Spectrofluorometer experiments with a Fluorolog-2 (Spex Industries, Inc. NJ). Both instruments were controlled from a 486-33SX computer (Zenith Data Systems) by CHI660 electrochemical workstation Version 1.02 and DM3000 version 3.32 software respectively.

The electrochemical cell was a spectrophotometric cuvette (12 mm x 12 mm) placed in the cell holder of the spectrofluorometer. The experimental set up for this cell is shown in Figure 1. W is the working electrode, R the reference electrode and A the auxiliary electrode. Working and auxiliary electrodes were prepared from platinum wire (99.9 %, 1.0 mm diameter, Aldrich) and the reference electrode was an Ag/AgCl electrode (CH Instruments). Fluorescence measurements were collected using a quartz fluorometric cuvette of the same dimensions.
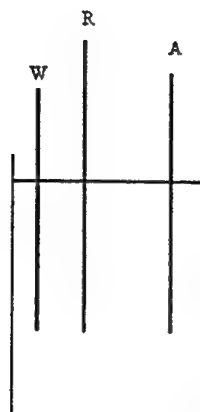
Figure 1 Schematic of the electrochemical cell set up.

The sample cell was placed in the sample compartment of the fluorometer for spectral measurements. All emission slits were removed to allow as much light as possible to pass through to the PMT. Stray light was excluded by using black plastic sheeting or black electrical insulation tape to cover openings where stray light could enter. A block diagram of the apparatus used for the ECL spectrum acquisition is shown in Figure 2.
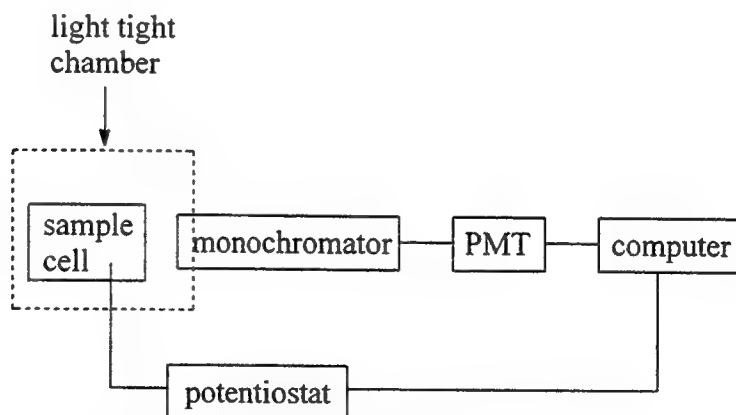
Figure 2 Block diagram of the experimental apparatus configuration.

Procedure

Solutions of the DATs, DNT and 3,4,5-T were scanned for fluorescence immediately after preparation. Scans which showed the most intense emission were then run on a daily basis. In order to determine if the changes in fluorescence correlated with changes in ECL emission, ECL scans were run on the DAT and 3,4,5-T at various intervals. Fluorescence scan were also run on the samples listed in Table 1 to determine if any of these samples gave rise to any significant fluorescence. Fluorescence scans were run with the emission monochromator set at 50 nM intervals from 400 to 700 nM and the excitation wavelength scanned from 250 nM to 30 nM below the emission monochromator setting. All slits on the fluorometer were 2.5 mm in width and the PMT voltage was set at 950 V.

Once satisfactory cyclic voltammagrams had been obtained for the standard solution of potassium ferricyanide (Figure 3), then a solution of $Ru(bpy)_3^{2+}$ (1.5 mL) and TPA (1.5 mL) was used as a test ECL solution. A potential of 1.2 V was applied to the solution and the cuvette inspected in a slightly darkened room for any indication of ECL.
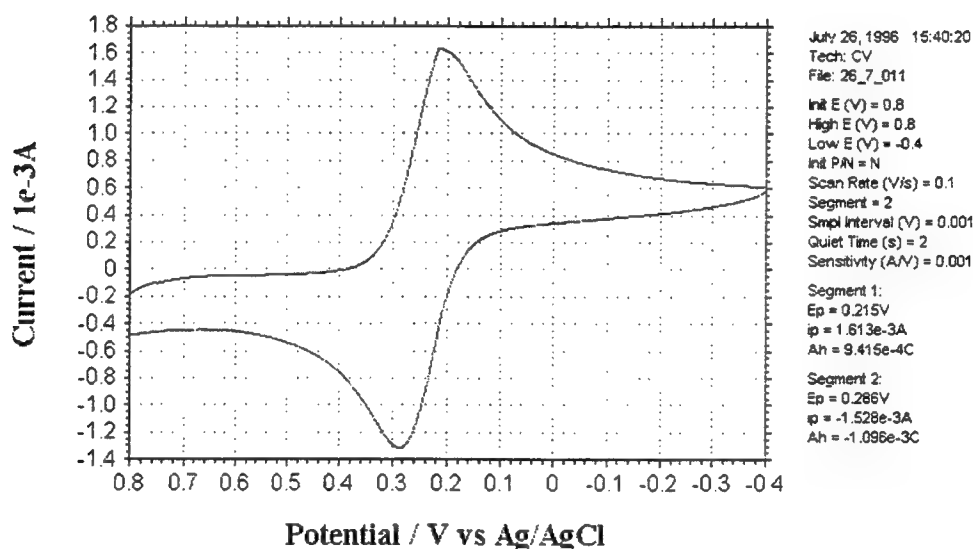
Figure 3 Cyclic voltammagram of the test potassium ferricyanide solution.

To determine the ECL spectra of other solutions a mixture of the sample, TPA and PBS (where necessary) was placed in the cuvette. This was then placed into the sample holder of the fluorometer, potential applied and the wavelengths scanned for emission. A blank file of PMT dark current (obtained with no potential applied) was automatically subtracted from the acquired data file. The PMT was cooled with a thermoelectric cooler (Products for Research, MA) and cooling water chilled to 15 °C with a circulating water chiller (Neslab, NH).

Results

Fluorescence of DATs, DNT and 3,4,5-T:

The solutions of DAT and 3,4,5-T showed strong fluorescence immediately after preparation. However the fluorescence of the DAT samples showed considerable change over time. Figures 4 shows the excitation scans of 2,4-DAT with the emission monochromator set at 400 (scan 1), 450 (scan 2) and 600 (scan 5) nM respectively obtained on the day of solution preparation (day 1). Figures 5 shows the excitation scans for 3,4-DAT under the same conditions on the same day.
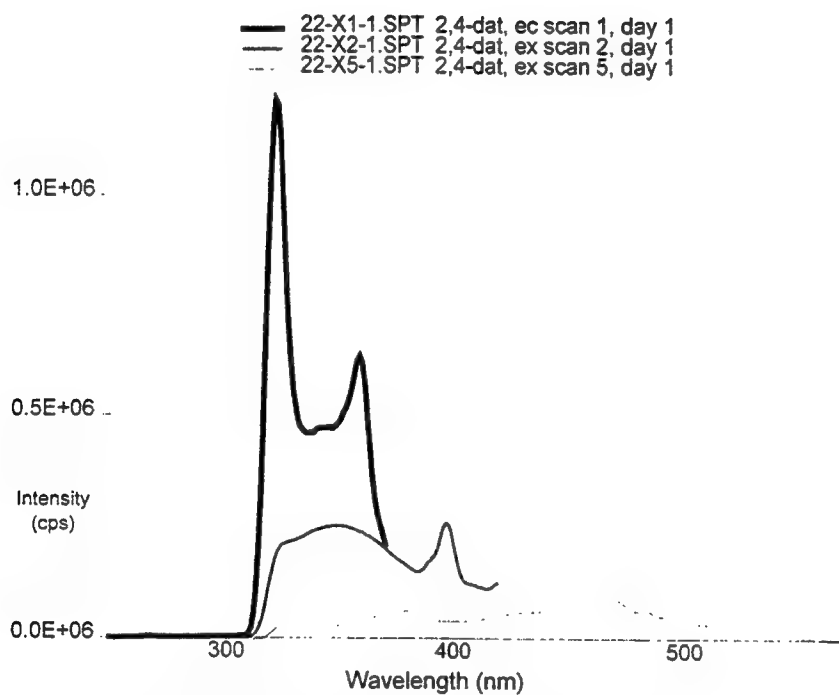
Figure 4 Fluorescence excitation scans of fresh 2,4-DAT in methanol.
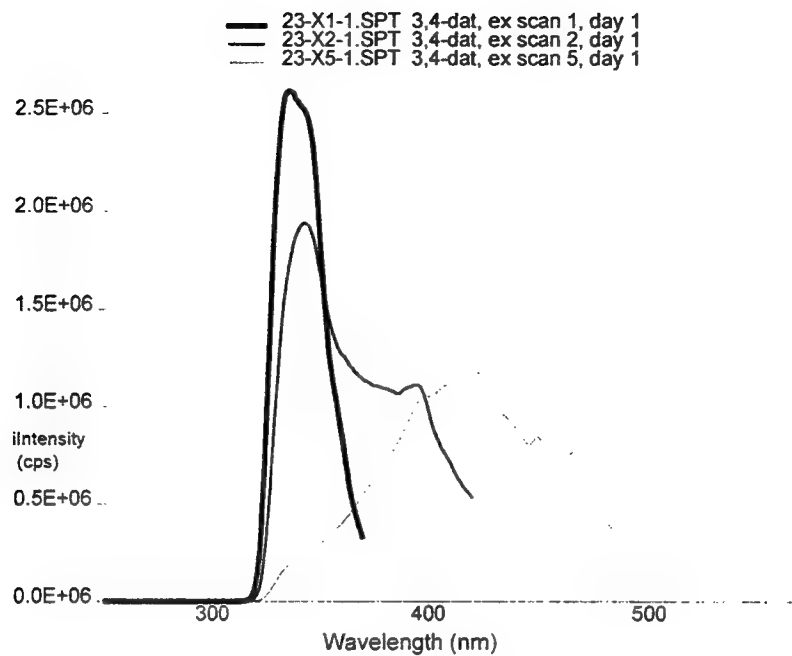
Figure 5 Fluorescence excitation scans of fresh 3,4-DAT in methanol.

Figures 6 and 7 illustrate these same scans on the same solutions 16 days after solution preparation. Note the intensity scale on Figure 7.
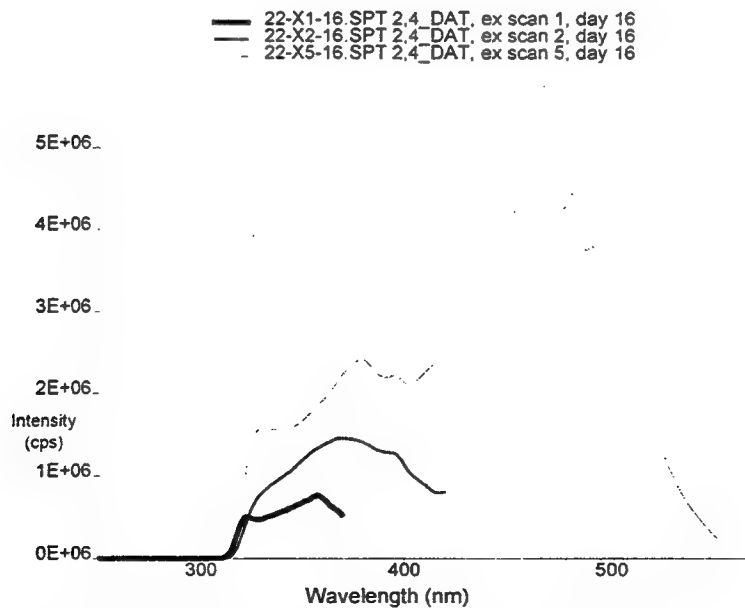
Figure 6 Fluorescence excitation scans of 2,4-DAT in methanol after 16 days.
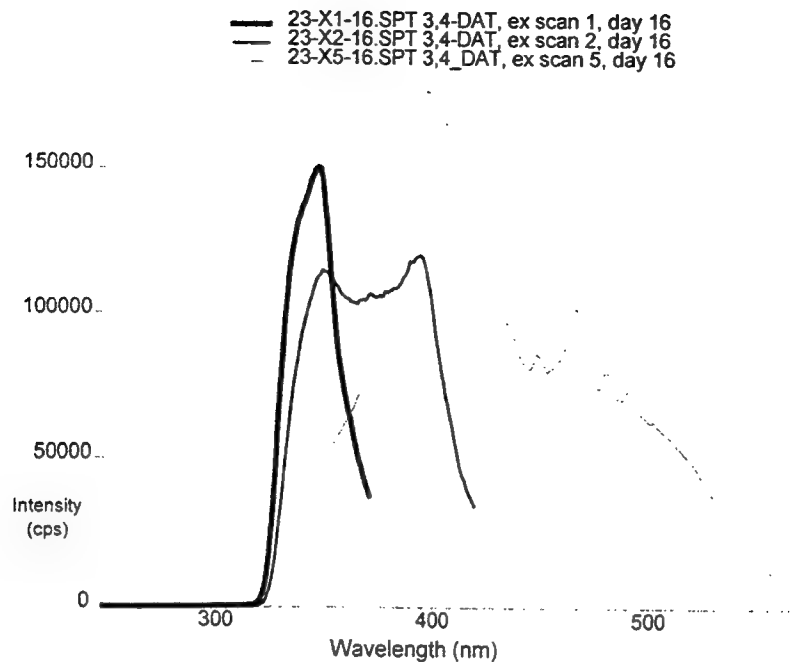
Figure 7 Fluorescence excitation scans of 3,4-DAT in methanol after 16 days.

The changes seen over time are illustrated in Figures 8-13 which show the excitation scans for 2,4-DAT and 3,4-DAT with emission at 400, 450 and 600 nM on days 1, 5, 9 and 15.
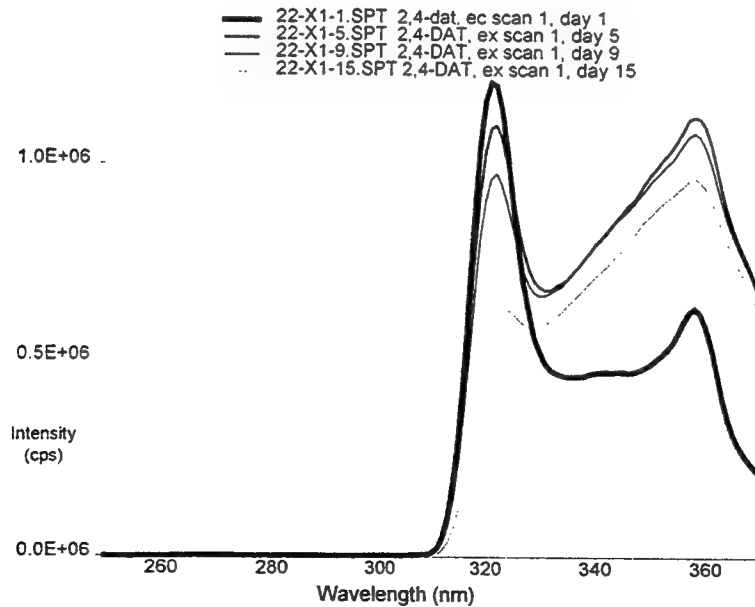
Figure 8 Fluorescence excitation scans of 2,4-DAT in methanol with emission at 400 nM.
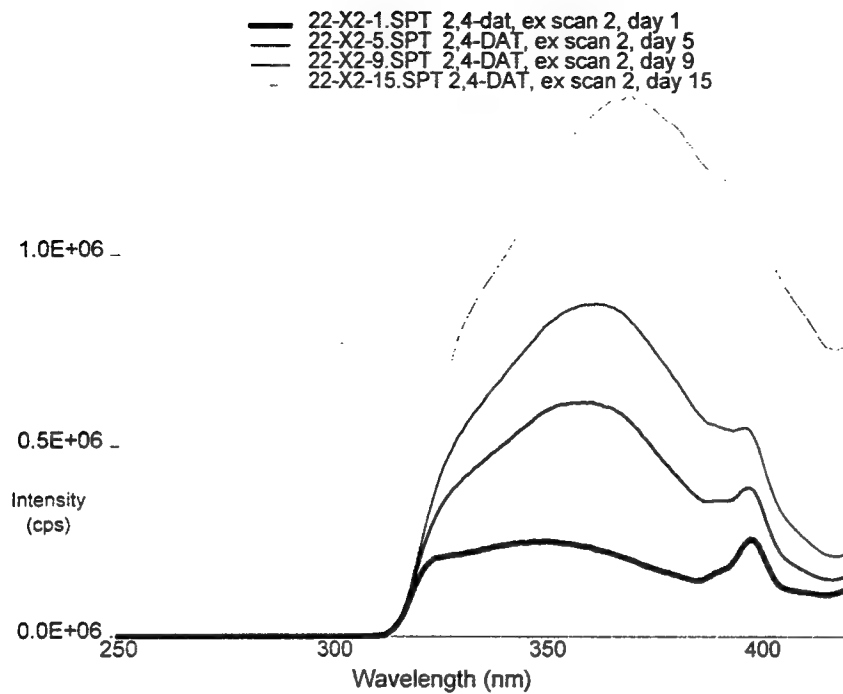
Figure 9 Fluorescence excitation scans of 2,4-DAT in methanol with emission at 450 nm.
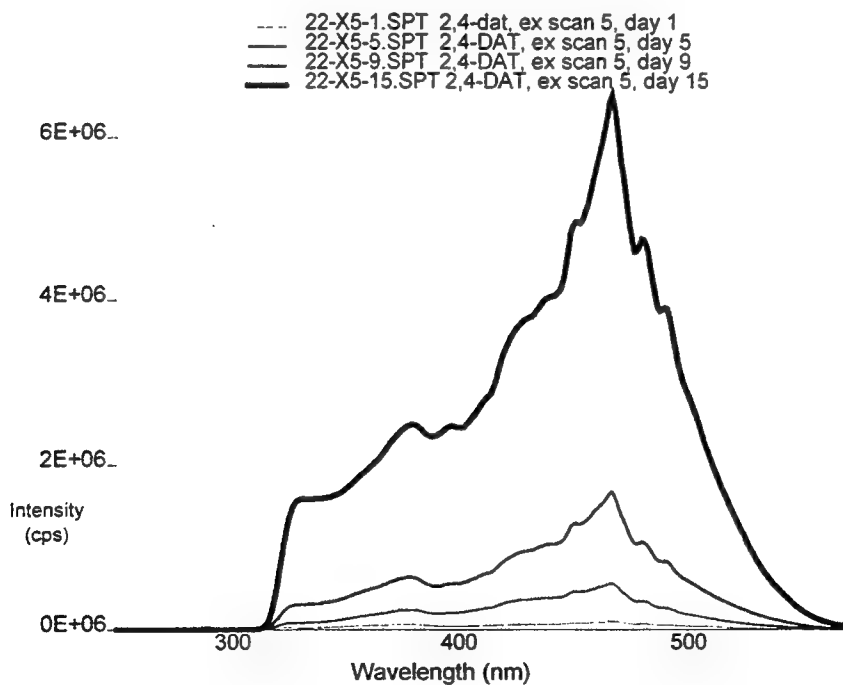
Figure 10 Fluorescence excitation scans of 2,4-DAT in methanol with emission at 600 nm.
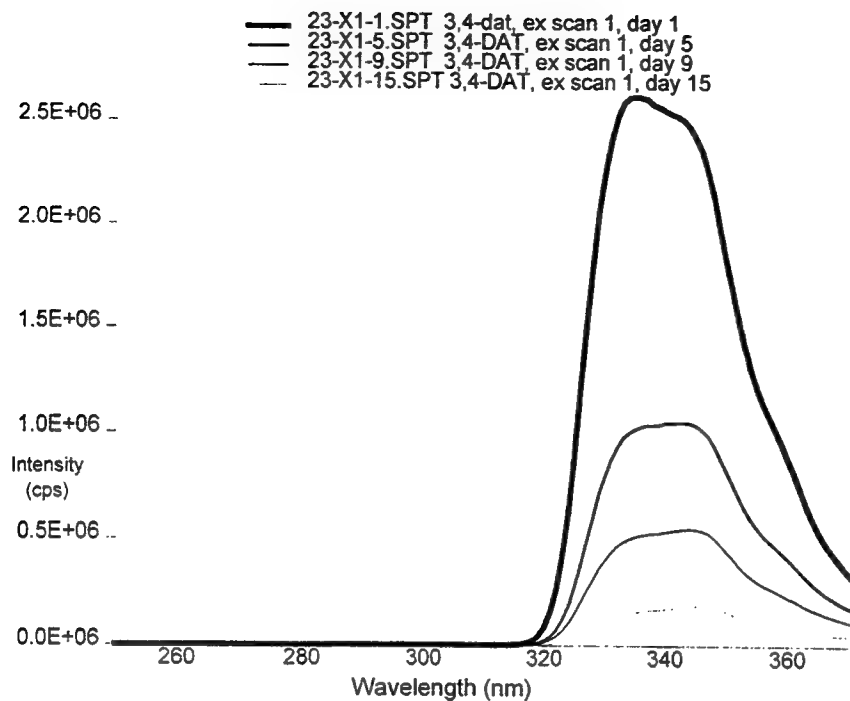
Figure 11 Fluorescence excitation scans of 3,4-DAT in methanol with emission at 400 nm.
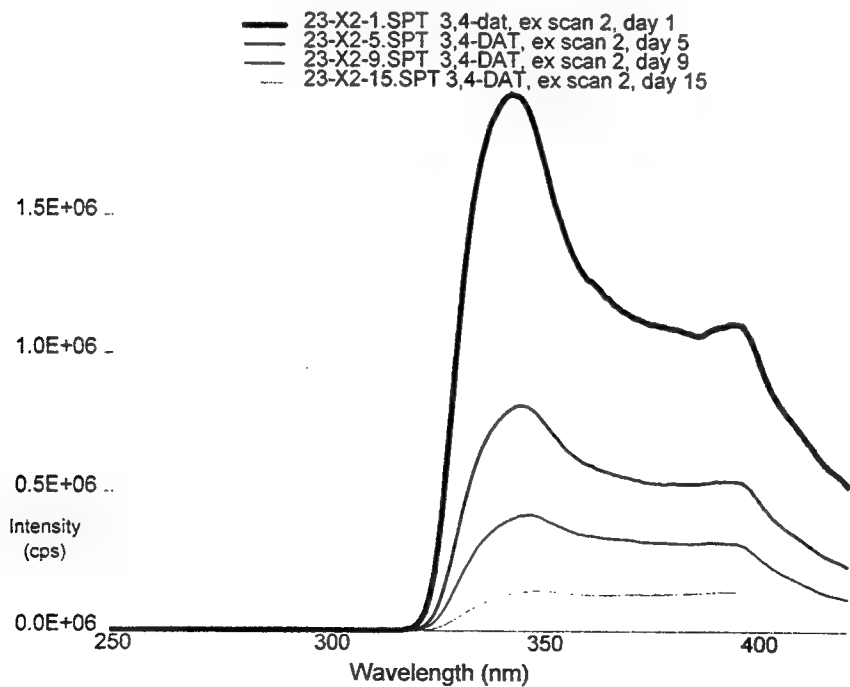
Figure 12 Fluorescence excitation scans of 3,4-DAT in methanol with emission at 450 nm.
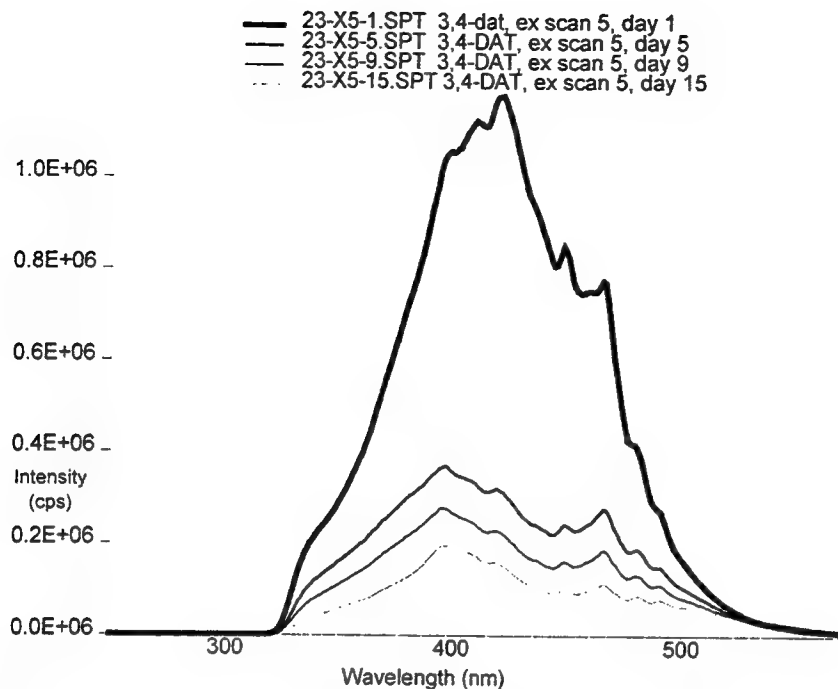
Figure 13 Fluorescence excitation scans of 3,4-DAT in methanol with emission at 600 nm.

The 3,4-DNT showed very slight emission at 550 nM immediately after preparation. This emission decreased slightly over the time monitored. No traces are shown here as the emission was small (< 250 000 cps), even with the fresh solution.

The 3,4,5-T showed two excitation maxima, with emission at 400 and 500 nM. This emission was much more stable than that seen from the DAT samples. The excitation spectrum for emission at 400 nM showed very little change over the same time period as the DAT solutions illustrated earlier, this is shown in Figure 14. The excitation spectrum for emission at 500 nM showed a slight decrease over this time period. The spectra for this are shown in Figure 15.
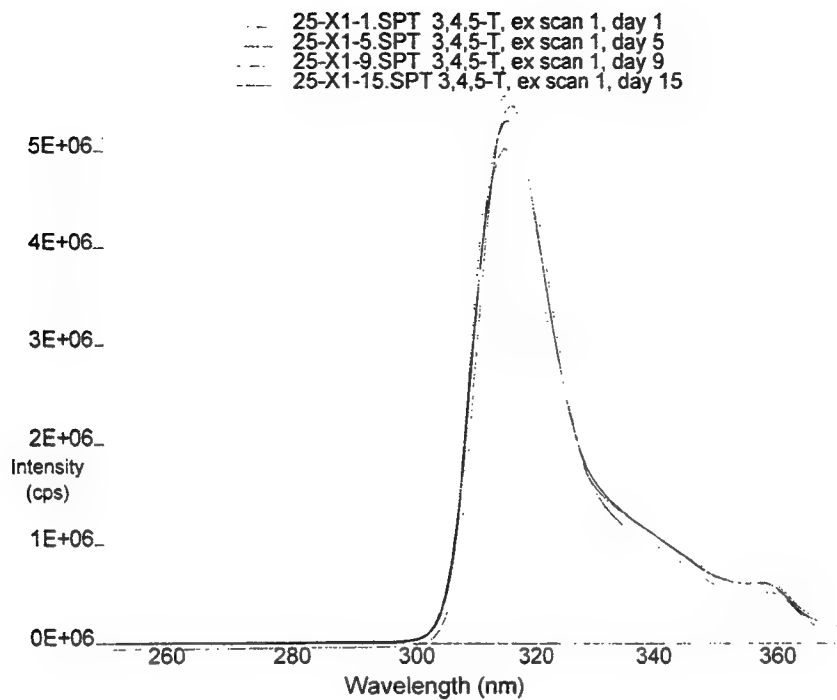
Figure 14 Fluorescence excitation scans of 3,4,5-T in methanol with emission at 400 nm.
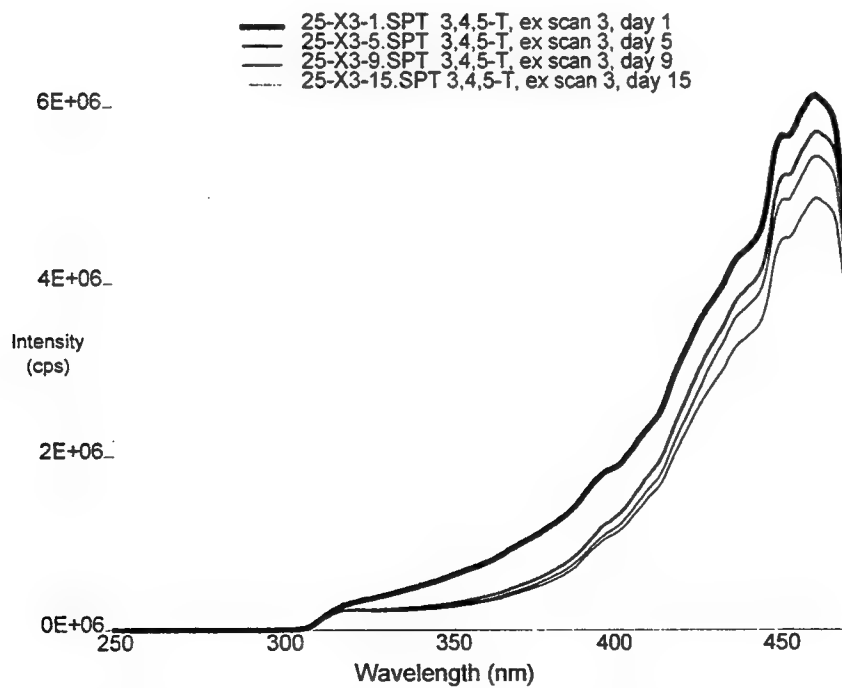
Figure 15 Fluorescence excitation scans of 3,4,5-T in methanol with emission at 500 nm.

2-14

Fluorescence of Other Samples:

Several of the samples studied showed fluorescence under the conditions used. Of particular interest was the fluorescence of the blood of *Molgula Occidentalis* (M. Occ.). The blood samples showed fluorescence with an excitation maxima of 372.0 nM and an emission maxima of 454.5 nM. Figure 16 shows an overlay of the excitation and emission spectra for sample 15. Interestingly the addition of protease K had no effect upon this fluorescence. All samples of M. Occ. body fluid examined showed the same fluorescence with there being no noticeable difference between serum, blood cell lysate and blood cells treated with protease K.

Preliminary experiments were conducted investigating the effect of metal ions on DAT fluorescence. Gold (198 μL of 1 mg/mL Spex reference standard) and copper (64 μL of 1 mg/mL Fisher reference standard) were added to 2,4-DAT and 3,4-DAT solutions (400 μL) respectively and the volume made up to 2 mL with methanol. The fluorescence of the solutions before and after the addition of metal ions was recorded, as well as a blank of methanol with the metal ion. An increase in emission and a shift to lower excitation wavelengths were seen in the excitation scans with emission at 400 and 450 nM for the 2,4-DAT/Au solution when compared to the 2,4-DAT solution. Similar results were obtained for the 3,4-DAT/Cu solution in comparison to the 3,4-DAT solution. An example of the changes seen is shown in Figure 17.

The fluorescence shown by samples 2 and 17 is inconclusive as the spectra were very different for these two samples. No more 2,3,4-T was available for fresh solution preparation to confirm whether 2,3,4-T has a similar excitation and emission spectra to 3,4,5-T. The DAT samples exhibited variance which was related to the date of solution preparation.
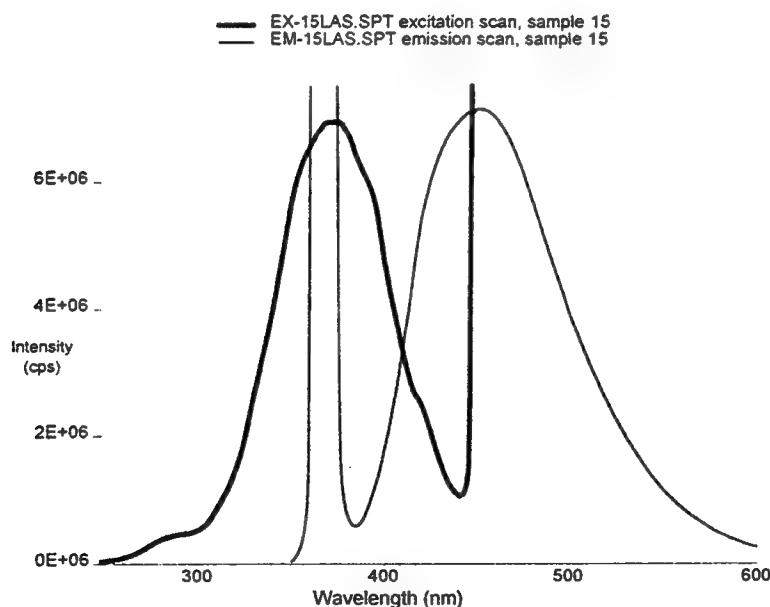
Figure 16 Fluorescence excitation (λ_{em} = 454.4 nM)and emission (λ_{ex} = 369 nM) spectra of sample 15.
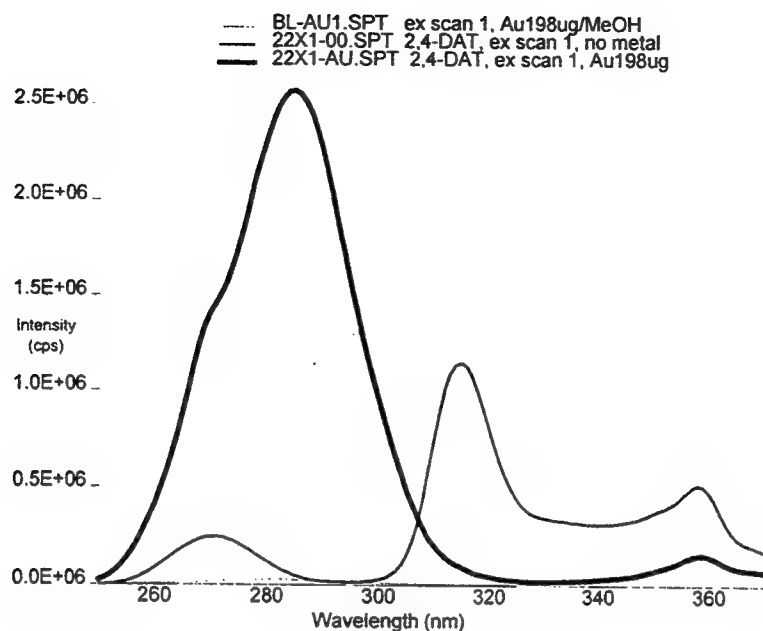
Figure 17 Excitation scan (λ_{em} = 400 nM) for 2,4-DAT with Au, 2,4-DAT alone and Au in methanol.

ECL of $Ru(bpy)_3^{2+}$:

The ECL generated by the $Ru(bpy)_3^{2+}$ /TPA solution was clearly visible, even in a room with fluorescent lights on except directly over the sample cuvette. A video of the ECL showing the immediate light production upon application of the potential and cessation of light emission upon termination of the applied potential was made and converted to digital form (11). A spectrum of this ECL emission was collected and is shown in Figure 18. The spectrum is corrected for PMT dark current, but not for PMT response. Maximum emission was seen at 604 nM. A comparison between this spectrum and the fluorescence emission (λ_{ex} = 529.5 nm) spectrum showed that they were identical. This is in line with previously published reports on the ECL spectrum of $Ru(bpy)_3^{2+}$ which give the emission maxima as 610 nM (1).

2-16
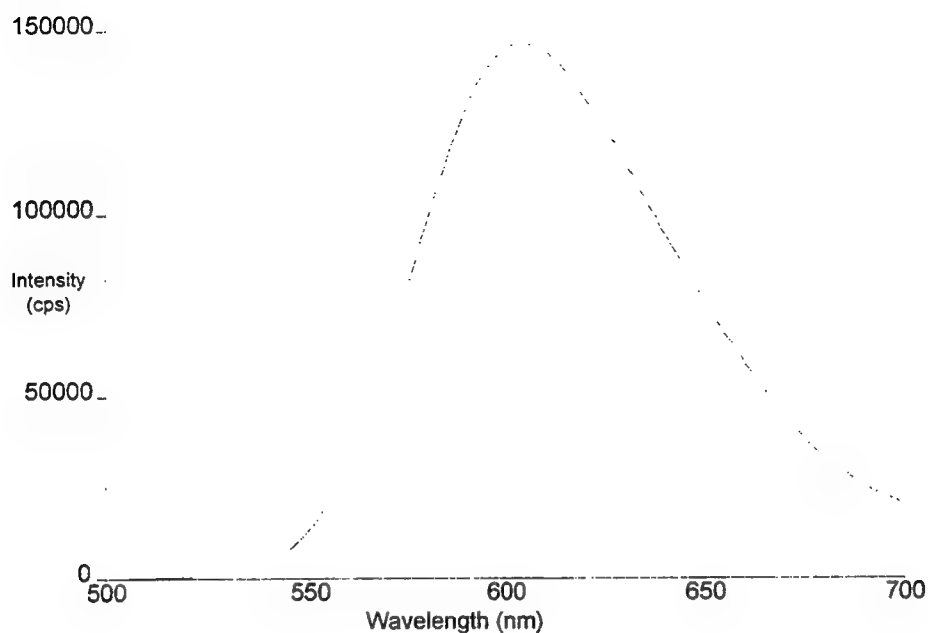
Figure 18 ECL spectrum of a $Ru(bpy)_3^{2+}$ /TPA solution with an applied potential of 1.2 V.

ECL of Other Samples:

Unfortunately no ECL spectra were obtained for any compounds other than $Ru(bpy)_3^{2+}$. Several methods of spectral acquisition were investigated. These include scanning the wavelengths whilst applying a potential to the solution of interest and holding at 50 nM wavelength intervals whilst integrating the signal for 10 s whilst applying the potential. Neither of these approaches yielded a signal above baseline. This was presumably due to the poor light throughput of the double monochromator system in the Spex fluorometer.

Discussion

The changes in fluorescence of the DAT solutions with time is of great interest. This suggests that the DATs are not inactive in solution but undergo some sort of transformation. This is not likely to be a transformation to a DNT as the test DNT solution showed no fluorescence. The changes appeared to have no detrimental affect upon the ECL seen. In fact as the solutions aged it appeared that ECL intensity increased. Further studies are needed to confirm this. This may be important in the determination of DAT in the environment, especially as related to explosive degradation products.

2-17

There is also evidence for interactions between the DATs and added metal ions. Previous work has speculated that the interactions are size dependent in relation to the metal ion (12). The investigation of other metal ions could confirm whether interactions that affect the fluorescence occur with all metal ions or just those that affect the ECL emission. A cyclic voltammagram from 0 to 1.5 V at 0.1 V/s with a sensitivity of 1×10^{-2} A also showed a change upon addition of Cu ions to a solution of 3,4-DAT and PBS. Confirmatory studies need to be conducted in this area and no results from this work are presented in this report.

A preliminary investigation of the ECL reaction potential for 3,4-DAT/Cu indicated that maximum emission was obtained at 2.4 V, with no emission seen below a potential of 1.6 V. A similar investigation of the M. Occ. blood ECL reaction gave 1.4 V as giving the maximum emission, with no emission occurring below 1.0 V. From these results it is not clear if TPA is even required for the ECL emission to occur. No experiments were possible in this area due to instrument restrictions.

A dedicated instrument specific for ECL will be required to acquire the spectra of DAT and M. Occ. blood ECL reactions. This is because the low light levels output by these reactions are insufficient to allow the use of make shift instrumentation as used in this project. The light throughput of the monochromator system, in conjunction with the sample cell alignment do not allow enough light to reach the PMT.

Conclusions

The results presented here clearly show a change in fluorescence characteristics over time for DAT solutions prepared in methanol. In addition preliminary evidence for binding or interaction of 2,4-DAT with gold ions and 3,4-DAT with copper ions was seen. Further study is needed to elucidate the exact nature of these interactions. The fluorescence of M. Occ. blood has been confirmed and determined to be unaffected by protein cleaving agents.

A major goal of this research, the determination of the ECL spectra of a number of compounds, was not achieved. This was due to the low light levels emitted by these reactions. However the groundwork for a number of exploratory experiments to further probe the ECL reactions of interest was laid.

Future Work

Some suggestions for experiments to further probe the ECL reactions of M. Occ. blood and the DATs are:
- Continue scans on methanol DAT solutions on a weekly basis.
- Look at DATs in aqueous solution.
- Run HPLC and MS on the DAT methanol solutions to determine what the transformation products are.
- Look at 2,4-DAT with gold (vary Au concentration) - use UV-Vis spectroscopy as well.
- Look at 3,4-DAT with copper (vary Cu concentration) - use UV-Vis spectroscopy as well.

- Use CV to monitor changes upon addition of metal ions to DATs.
- Vary the concentration of the metal ions to look for trends/changes.
- Look at other metal ions as well to determine if they show similar effects.
- Acquire some 2,3,4-T to compare with 3,4,5-T.
- Run fluorescence and IGEN ECL scans on SEC fractions of Molgula blood.
- Integrate the ECL *Molgula* studies with the tunicate metal ion studies.

Future work should also focus on the construction of a dedicated ECL instrument which is smaller and simpler than the make shift instrumentation used for this work. A smaller instrument designed for ECL work would be more sensitive than the instrumentation used here and allow for the determination of the low light levels generated by the ECL reactions of interest.

An instrument specific for ECL would not only release the two instruments used previously for more conventional use but also allow more detailed study of ECL reactions. This instrument should be capable of making ECL spectral measurements in both static and flowing solutions with the ability to vary both the voltage and the scan rate.

The advantage to be gained from further study into biological ECL is the potential for sensitive and selective biosensors using the reactions characterised by the instrument. These biosensors or other ECL sensors could be used for remote automated monitoring of metal ions or organic analytes. The potential for a portable, fieldable instrument which could be use by unskilled personnel also exists.

All these potential applications require successful characterisation of ECL reactions. The experience gained in constructing an ECL instrument to monitor these reactions will be invaluable when applied to development of a fieldable instrument.

References

1)   A.W. Knight and G.W. Greenway, *Analyst (London)*, 1994, 119, 879.

2)   J.B. Noffsinger and N.D. Danielson, *J. Chromatogr.*, 1987, 387, 520.

3)   N.D. Danielson, L. He, J.B. Noffsinger and L. Trelli, *J. Pharm. Biomed. Anal.*, 1989, 7, 1281.

4)   S.N. Brune and D.R. Bobbitt, *Talanta*, 1991, 38, 419.

5)   T.M. Downey and T.A. Nieman, *Anal. Chem.*, 1992, 64, 261.

6)   D.E. Ryan, N.D. Ghatlia, A.E. McDermott, N.J. Turro, K. Nakanishi, and K. Kustin, *J. Am. Chem. Soc.*, 1992, 114, 9659.

7)   J.G. Bruno, S.B. Collard, D.J. Kuch, and J.C. Cornette, *J. Biolumin. Chemilumin.*, In Press.

8)   P. T. Kissinger and W. R. Heineman, *J. Chem. Educ.*, 1983, 60, 702.

9)   P. T. Kissinger and W. R. Heineman, *J. Chem. Educ.*, 1983, 60, 772.

10)   E. Kime-Hunt, K. Spartalian, S. Holmes, M. Mohan, and C.J. Carrano, *J. Inorg. Biochem.*, 1991, 41, 125.

11)   An MPEG version of this video is available on the World Wide Web at:
      http://main.chem.ohiou.edu/~andrews/andrews.html.

12)   J.G. Bruno and J.C. Cornette, *Microchem. J.*, In Press.

DEVELOPMENT OF A PRIMARY CELL CULTURE PREPARATION FOR STUDYING MECHANISMS GOVERNING CIRCADIAN RHYTHMICITY IN SUPRACHIASMATIC NEURONS.

Deborah L. Armstrong
Professor
Division of Life Sciences

The University of Texas at San Antonio
6900 Loop 1604 N.W.
San Antonio, Texas 78249

DEVELOPMENT OF A PRIMARY CELL CULTURE PREPARATION FOR STUDYING MECHANISMS GOVERNING CIRCADIAN RHYTHMICITY IN SUPRACHIASMATIC NEURONS.

Deborah L. Armstrong
Professor
Division of Life Sciences
The University of Texas at San Antonio

Abstract

The aim of the present study was to establish a cell culture system for characterizing the physiological and biochemical processes by which neuronal activity of the suprachiasmatic nucleus (SCN) is regulated. Whole-cell patch clamp recording experiments were conducted in both hippocampal and hypothalamic neurons in primary cultures from embryonic and postnatal tissue, respectively. Morphological examinations revealed a much more heterogeneous population of neurons in hypothalamic cultures and a lower percentage of cell survival. Tight membrane seals for stable recordings were difficult to achieve. However, spontaneous and evoked action potentials were observed indicating that voltage-gated ionic channels were functional in hypothalamic neurons culture day 7 and 14 when recordings were carried out. These results provide a foundation for our on going experiments to optimize the health and survival of SCN neurons in hypothalamic cultures.

DEVELOPMENT OF A PRIMARY CELL CULTURE PREPARATION FOR STUDYING MECHANISMS GOVERNING CIRCADIAN RHYTHMICITY IN SUPRACHIASMATIC NEURONS.

Deborah L. Armstrong

INTRODUCTION

Circadian rhythmicity in mammals appears to be controlled by the action of a light-entrainable pacemaker located in the suprachiasmatic nuclei (SCN) of the basal hypothalamus (15,24). Bilateral destruction or surgical isolation of the SCN abolishes circadian rhythmicity in rodents (10,24), and transplantation of fetal SCN tissue into the third ventricle of SCN-lesioned hosts restores rhythmicity (5,13,25). Furthermore, circadian rhythms in single unit activity (7,8), 2-deoxyglucose utilization (17) and vasopressin release (6,7) persist in the isolated SCN *in vitro* indicating that rhythm generation is intrinsic to the nucleus. In the absence of temporal environmental cues, this pacemaker continues to measure time by an endogenous mechanism (clock), driving biochemical, physiological and behavioral rhythms that reflect the natural period of the pacemaker oscillation. This endogenous period is usually slightly different than 24 hours (i.e., circadian). When mammals are maintained under a 24 hour light-dark (LD) cycle, the circadian pacemaker becomes entrained such that the period of the pacemaker matches that of the LD cycle. Photic information is conveyed to the pacemaker via monosynaptic retinal afferents projecting along the retinohypothalamic tract (RHT) and terminating within the SCN (11,12,16,18,26) In the entrained state, pacemaker-driven circadian rhythms are phase-locked with the LD cycle. Photic entrainment of circadian rhythms occurs as a consequence of the phase specific effects of environmental light on the activity of the circadian pacemaker as defined by the phase response curve to light pulses administered to animals maintained under constant darkness (4,28). In nocturnal rodents, light pulses

administered during the early subjective night cause phase delays of the pacemaker while pulses delivered during the latter half of the subjective night cause phase advances (4,28). Light pulses given during the middle of the subjective day do not cause phase shifts. Light-induced shifts represent long-term alterations in pacemaker function. Although the RHT projection appears to be both necessary and sufficient to support photic entrainment of the pacemaker, little is known concerning the identity of SCN neurons responsive to photic stimulation or the neurochemical sequelae responsible for light-induced alterations of pacemaker activity.

Characterization of the mechanisms regulating SCN function will have important applications to human performance since it is known that a sudden phase shift of the LD cycle, such as that which occurs during transmeridian jet travel, results in internal desynchronization of physiological rhythms leading to a temporary period of malaise and suboptimal performance (jet lag). A similar condition may result when an imposed rest-activity cycle is incompatible with the LD cycle, such as occurs during shift work. An understanding of the neurochemical mechanisms responsible for the generation and photic entrainment of circadian pacemaker oscillations would facilitate the development of pharmacological strategies for altering clock function and, ultimately, alleviating the physiological and performance decrements associated with activities that are incompatible with the circadian system in man.

Progress over the last ten years have produced substantial insight into the neurochemical regulation of the mammalian circadian system. For example , it has been demonstrated that systemic administration of serotonergic compounds attenuate light-induced phase shifts on free-running activity in rodents and that both serotonin and the selective 5-HT_{1A} agonist 8-OH-DPAT, inhibit field potentials recorded in the SCN in response to optic

3-4

nerve stimulation using a hypothalamic slice preparation (23). In order to fully understand the underlying mechanisms responsible for regulation of the circadian system, technology must be developed to allow for the study of the circadian clock at the single cell level. The purpose of the experiments presented here is to establish an SCN cell culture system for characterizing the physiological and biochemical processes by which SCN neuronal activity is regulated. Since hippocampal neurons in culture provide a relatively hearty, homogeneous preparation for conducting patch-clamp recording experiments, we began our studies using this preparation. Information gained from these experiments was used to refine our procedures for identifying healthy SCN neurons *in vitro*.

METHODS

Experimental Animals. For both hippocampal and hypothalamic preparations pregnant Sprague Dawley rats were used and delivered every week such that the arrival date coincided with gestation day 18. Hippocampal tissue was harvested from fetuses on gestation day 19 and hypothalamic cultures were prepared from 4 to 6 day old pups.

Cell Culture Procedures. Hippocampal cultures were prepared using a procedure similar to that described by Goslin and Banker (8). The dam was anesthetized with Nembutal (60 mg/kg) and the fetuses removed and placed in chilled, sterile buffered saline. Hippocampi were removed and incubated for 15 min at room temperature in Hank's balanced salt solution (Ca^{++},Mg^{++} -free) containing 0.25% trypsin. The hippocampi were then passed 10-15 times through a small bore, fire-polished Pasteur pipette and the resulting dissociated cells collected by centrifugation and resuspended in plating medium consisting Eagle's MEM with Earle's salts containing extra glucose (600 mg/ml) and 10 % horse serum. The cells were plated into 35 mm Falcon microbiology dishes containing glass coverslips pretreated with poly-L-lysine and placed in an 36°C incubator containing a 5% CO_2 atmosphere. After 3 to

4 hours the coverslips were moved to dishes containing maintenance medium containing 5% horse serum, 100 µg/ml streptomycin, 100 units/ml penicillin, sodium pyruvate, ovalbumin and N-2 supplements (insulin, transferrin, putrescine, progesterone, and selenium dioxide). After 5 to 6 days 1 µM cytosine arabinoside was added to the medium to suppress glial proliferation.

Hypothalamic cultures were prepared using a procedure similar to that described by Baughman and colleagues (1). Pups were anesthetized with 100 mg/kg ketamine and sacrificed by cervical dislocation. Cultures were made from either punches of the SCN area or from dissections of a more extensive area that included the ventral medial hypothalamus. Tissue was enzymatically treated with papain for 30 min at 30° to 32°C and washed extensively to inhibit enzyme activity. The solution was centrifuged and the tissue resuspended in a triturating solution. After complete dissociation using fire polished pipettes the cells were plated at a density of 100,000 to 500,000 cells per ml. Plating and maintenance media were identical to those used for hippocampal neurons.

Recording Procedures. Currents were recorded at room temperature using an Axopatch 1D amplifier interfaced to an IBM 486 computer via a TL-1/Lab Master DMA acquisition system supported by pClamp version 5.5 software. Patch clamp recording pipettes were pulled from TW150F-4 glass using a Brown-Flaming model P-88 puller and a three step program. After fire polishing the tip resistances ranged from 2-3 M ohm. The bath and electrode solutions for isolating voltage-gated potassium currents were as follows (in mM): Bath - 137 NMDG, 5.4 KCl, 1 $MgCl_2$, 2 $CaCl_2$, 10 Glucose, 10 HEPES (pH 7.4); Electrode - 145 KCl, 1 $MgCl_2$, 2.2 EGTA, 5 HEPES (pH 7.2). For recording I_{ca} generated by voltage-gated calcium channels we used: Bath - 139 TEA-Cl, 5 4-AP, 5 $CaCl_2$, 15 HEPES, 15 Glucose (pH 7.4); Electrode - 124 CsCl, 0.15 KCl, 1.0 $MgCl_2$, 2.2 EGTA, 10 HEPES (pH 7.2). Na_2-ATP (3 mM) was added to

all pipette solutions. Positive pressure was applied to the pipette while entering the bath and approaching the cell then negative pressure to produce a seal. Formation of a usable seal was monitored and seals less than 1 G ohm in resistance were rejected. Additional negative pressure was applied to rupture the cell membrane for whole-cell patch clamp recording. Series resistance and cell capacitance were calculated. For sodium and potassium currents the membrane potential was stepped between -100 and +40 mV in 10 mV increments from a -70 mV resting potential. Three trials were averaged and current measurements were made at the plateau phase of the current. To measure voltage-gated calcium currents cells were held at -80 mv and stepped to 0 mV for 30 ms and the peak current.

RESULTS

The morphological characteristics of cells in our hippocampal preparations correspond with those described by Goslin and Banker (8). These neurons are derived mainly from areas CA1 and CA3 regions of the fetal hippocampus at this stage of development. The pyramidal cells begin to extend neurites within hours of attachment to the coverslip surface and within 5 days extensive neurite outgrowth is apparent (Figures 1 and 2). The morphology of neurons derived from the hypothalamus is much more diverse and the cell density is much lower (Figure 3).

All of the hippocampal neurons tested (n=11) displayed robust voltage-gated potassium currents; however, only 5 out 8 displayed stable sodium currents in response to a full range of membrane depolarizing steps. The majority of the recording experiments were conducted on cell culture day 7 to 10 day and it is possible that voltage-gated sodium channels are not yet fully expressed at this time. Examples, of whole-cell current traces are shown in Figure 4. Stable patch-clamp recording from hypothalamic neurons proved to be very difficult and

the majority of responses obtained were recorded under current clamp mode and consisted
of spontaneous and evoked action potentials (Figure 5).
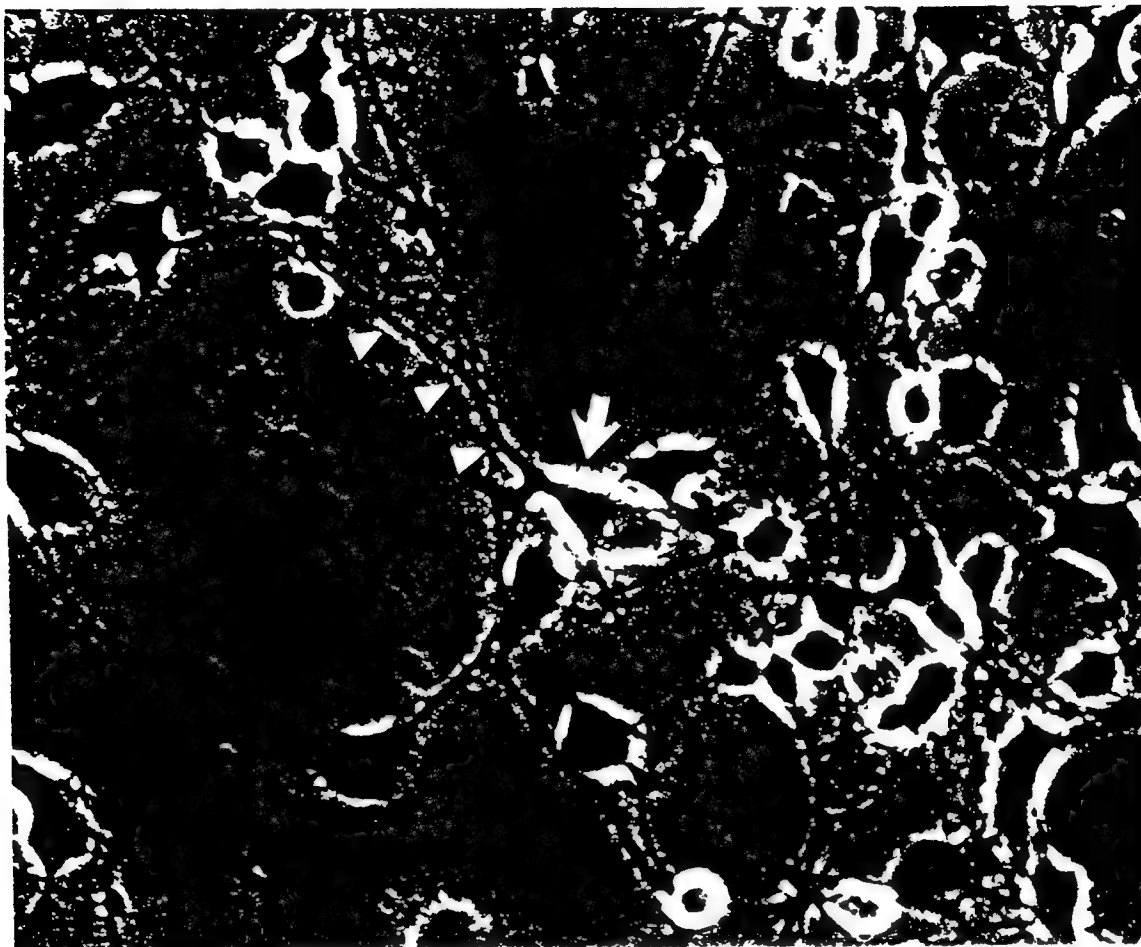
Figure 1. Phase-contrast micrograph of hippocampal neurons after 5 days in culture. The hippocampi were dissected from rat fetuses at gestation day 18. At this developmental stage, the granule cells of the dentate gyrus have not yet developed; therefore, the culture consists primarily of pyramidal cells with a small number of interneurons. The cell body of the neuron highlighted with the large white arrow displays the typical triangular shape of neurons forming the hippocampal stratum pyramidal layer. A primary apical dendrite can also be seen extending some distance from the soma (small arrowheads). With phase-contrast optics, healthy neurons exhibit bright cell bodies.
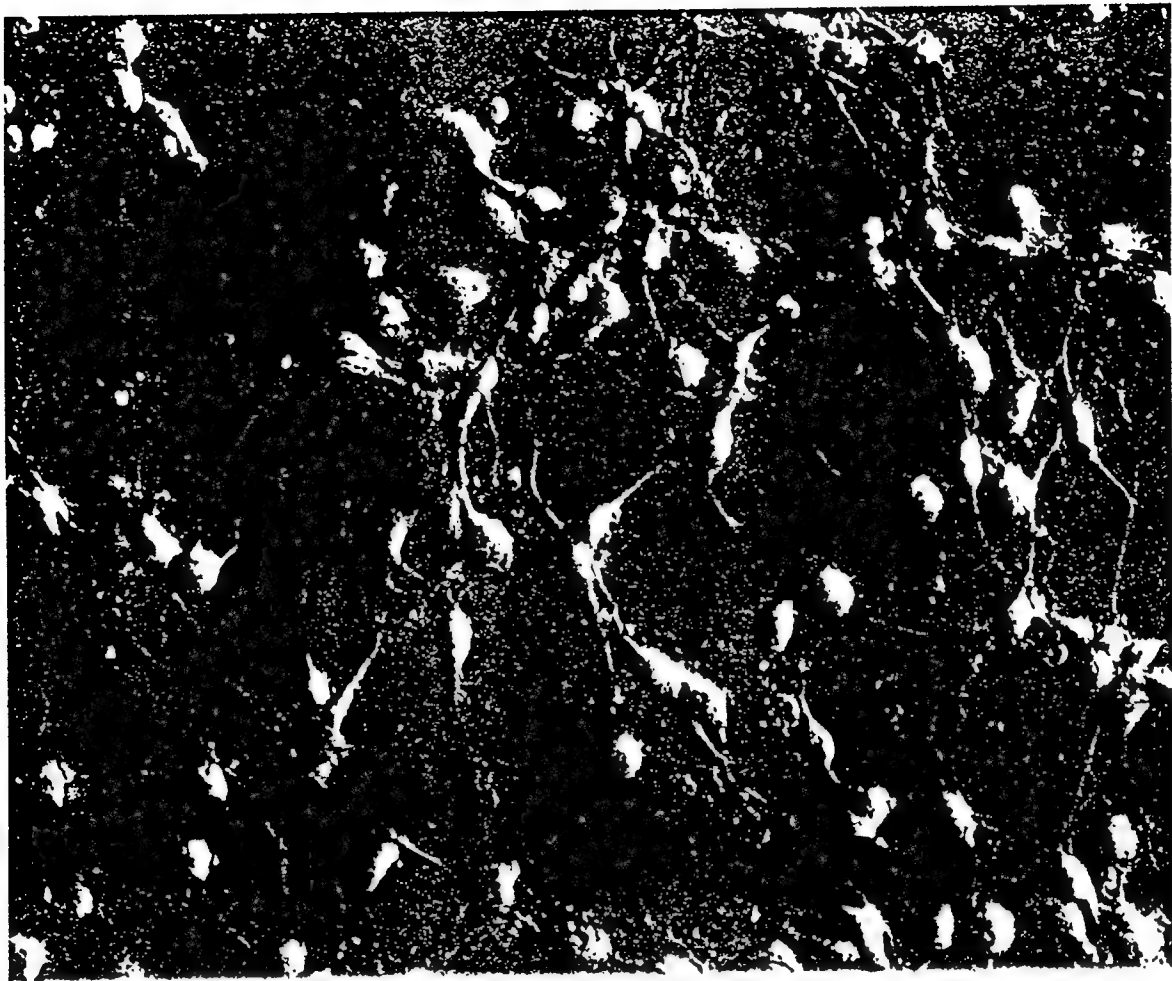
Figure 2. Micrograph of hippocampal neurons after 5 days in culture using Hoffman Modulation optics. Hoffman Modulation produces a three dimensional image which facilitates electrode placement on the domed surface of the soma. Under these observation conditions, healthy neurons are those displaying a smooth plasma membrane surface.
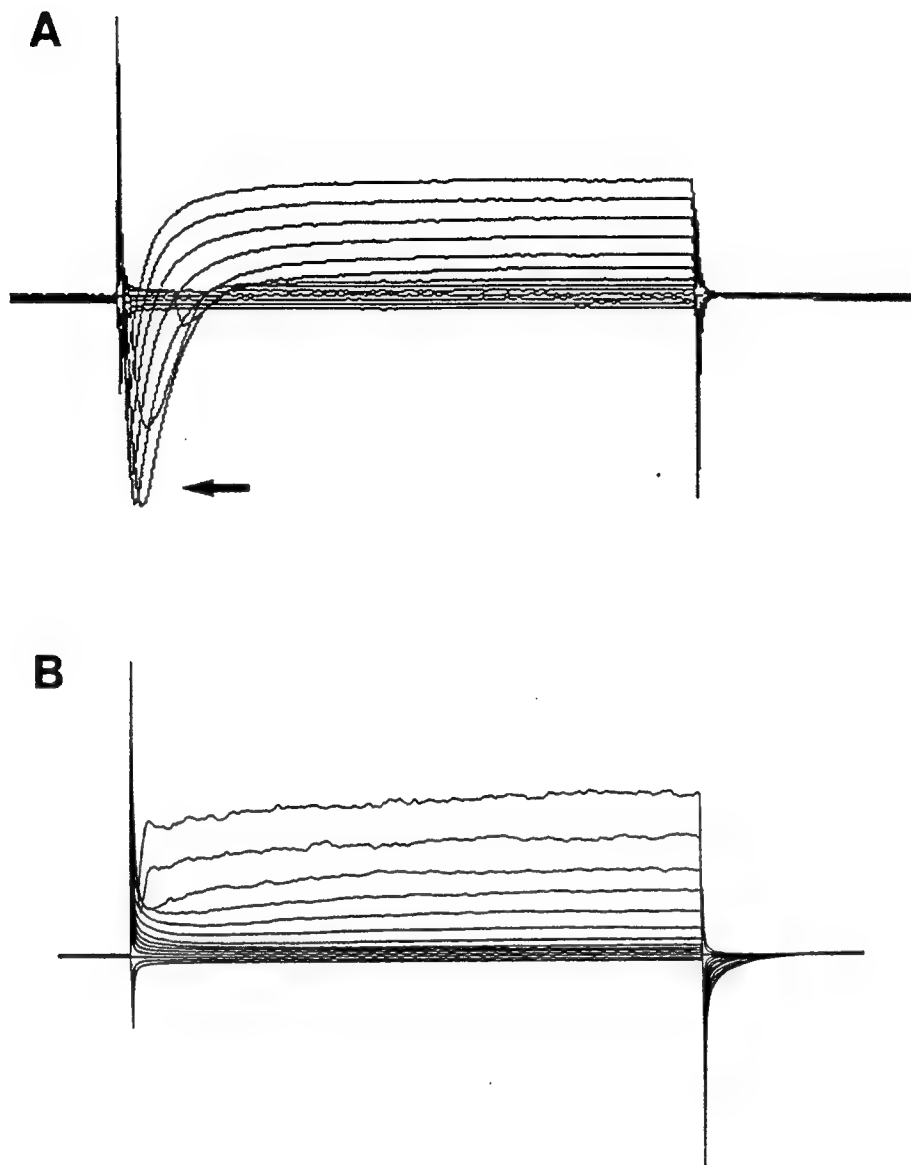
Figure 4. Whole cell currents generated by voltage-gated channels in cultured hippocampal neurons. A) Total current generated from -50 to +20 voltage steps in 10 mV increments from a -60 mV resting potential. The sharp inward currents (arrow) are generated by increased sodium and calcium ion conductance. Traces are the average of three trials. B) Isolated potassium currents recorded when extracellular N-methyl-D-glucamine (an ion impermeate to voltage-gated sodium channels) is substituted for sodium. Membrane potential stepped from -100 to +40 mV. Average of three trials shown; resting potential -70 mV and leak current is subtracted.

3-10
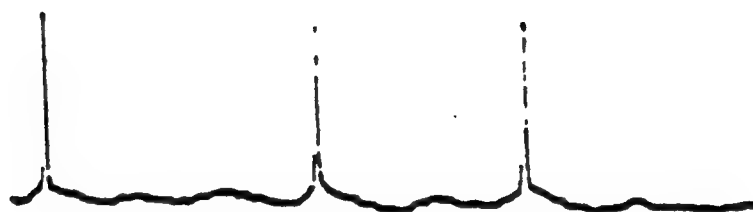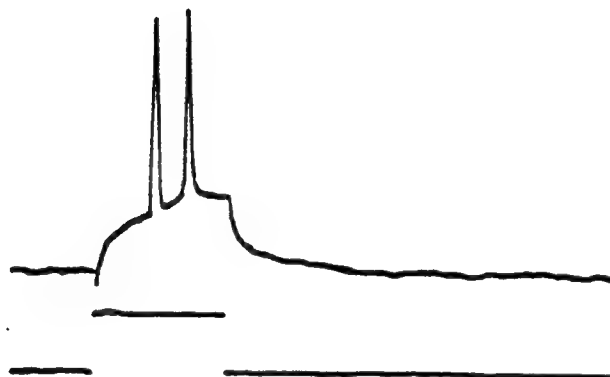
Figure 5. Responses of two hypothalamic neurons recorded in current-clamp mode shortly after seal formation. A) Spontaneous action potentials observed in a neuron after 9 days in culture. The resting membrane potential was -61 mV. B) Action potential evoked in neuron by a 1.0 nA positive current step. The resting membrane potential of the cell was -67 mV.

3-11

DISCUSSION

The results of these experiments confirm that our laboratory has successfully established the protocols and techniques needed for studying neurons in primary cultures. Although we are disappointed in the poor survival of hypothalamic neurons several possible solutions to this problem are being explored. The first approach will be to increase the cell plating density. Although we are adding conditioned medium obtained from glial cultures, it is possible that the neurons themselves are secreting nutrient factors which are not available in high enough concentrations under our present culture conditions. In addition, we observed that many hypothalamic culture dishes were contaminated with fibroblasts which could create unfavorable conditions for neuronal maturation. The source of these fibroblasts is most likely the remnants of meninges left on the ventral surface of the brain prior to dissecting out the hypothalamic tissue. Greater care during the dissection will resolve this problem.

The overall goal of this research is the elucidation of the mechanisms by which neurotransmitters such as serotonin regulate the mammalian circadian clock. We feel confident that the hypothalamic cell culture preparation we are using, which includes cells from the SCN, will provide a useful tool for obtaining this goal.

REFERENCES

1. Baughman, R.W.; Huettner, J.E.; Jones, K.A.; Khan, A.A. (1991) Cell culture of neocortex and basal forebrain from postnatal rats. In: *Culturing Nerve Cells*, eds. G. Banker and G. Goslin; Massachusetts, MIT Press, pgs. 227-249.

2.   Cahill, G.M.; Menaker, M. (1987) Kynurenic acid blocks suprachiasmatic nucleus responses to optic nerve stimulation. *Brain Res.* 410:125-129.

3.   Cahill, G.M.; Menaker, M. (1989) Effects of excitatory amino acid receptor antagonists and agonists on suprachiasmatic nucleus responses to retinohypothalamic tract volleys. *Brain Res.* 479: 76-82.

4.   Daan, S.; Pittendrigh, C.S. (1976) A functional analysis of circadian pacemakers in nocturnal rodents: II. The variability of phase response curves. *J. Comp. Physiol.* 106:253-266.

5.   DeCoursey, P.J.; Buggy, J. (1989) Circadian rhythnicity after neural transplant to hamster third ventricle: Specificity of suprachiasmatic nuclei. *Brain Res.* 500: 263-275.

6.   Earnest, D.J.; Sladek, C.D. (1986)   Circadian rhythms of vasopressin release from individual rat suprachiasmatic explants in vitro. *Brain Res.* 382:129-133.

7.   Gillette, M.U.; Reppert, S.M. (1987) The hypothalamic suprachiasmatic nuclei: Circadian patterns of vasopressin secretion and neuronal activity in vitro. *Brain Res. Bull.* 19: 135-139.

8.   Goslin, K.; Banker, G. (1991) Rat hippocampal neurons in low-density cultures. In: *Culturing Nerve Cells*, eds. G. Banker and G. Goslin; Massachusetts, MIT Press, pgs. 251-281.

9.   Green, D. J.; Gillette, R. (1982) Circadian rhythm of firing rate recorded from single cells in the rat suprachiasmatic brain slice. *Brain Res.* 245:198-200.

10.   Inouye, S. I. T.; Kawamura, H. (1979) Persistence of circadian rhythmicity in a mammalian hypothalamic "island" containing the suprachiasmatic nucleus. *Proc. Natl. Acad. Sci. USA* 76:5962-5966.

11.   Johnson, R. F.; Morin, L. P.; Moore, R. Y. (1988a) Retinohypothalamic projections in the hamster and rat demonstrated using cholera toxin. *Brain Res.* 462:301-312.

12.   Johnson, R. F.; Moore, R. Y.; Morin, L. P. (1988b) Loss of entrainment and anatomical plasticity after lesions of the hamster retinohypothalamic tract. *Brain Res.* 460:297-313.

13.   Lehman, M. N.; Silver, R.; Gladstone, W. R.; Kahn, R. M.; Gibson, M.; Bittman, E. L. (1987) Circadian rhythmicity restored by neural transplant. Immunocytochemical characterization with the host brain. *J. Neurosci.* 7:1626-1638.

14.   Liou, S. Y.; Shibata, S.; Iwasaki, K.; Ueki, S. (1986) Optic nerve stimulation induced increase of release of 3H-glutamate and 3H-aspartate but not 3H-GABA from suprachiasmatic nucleus in slices of rat hypothalamus. *Brain Res. Bull.* 16:527-531.

15.   Meijer, J. H.; Rietveld, W. J. (1989) Neurophysiology of the suprachiasmatic circadian pacemaker in rodents. *Physiol. Rev.* 69:671-707.

16.   Moore, R. Y.; Lenn, N. J. (1971) A retinohypothalamic projection in the rat. *J. Comp. Neurol*. 146:1-14.

17.   Newman, G. C.;   Hospod, F. E. (1986) Rhythm of suprachiasmatic nucleus 2-deoxyglucose uptake in vitro. *Brain Res*. 381:345-350.

18.   Pickard, G. E. (1982) The afferent connections of the suprachiasmatic nucleus of the golden hamster with emphasis on the retinohypothalamic projection. *J. Comp. Neurol*. 211: 65-83.

19.   Ralph, M. R.; Foster, R. G.; Davis, F. C.; Menacker, M. (1990) Transplanted suprachiasmatic nucleus determines circadian period. *Science* 247: 975-979.

20.   Ralph, M. R.; Menaker, M. (1985) Bicuculline blocks circadian phase delays but not advances. *Brain Res*. 325: 362-365.

21.   Ralph, M. R.; Menaker, M. (1986) Effects of diazepam on circadian phase and delays. *Brain Res*. 372: 405-408.

22.   Ralph, M. R.; Menaker, M. (1989) GABA regulation of circadian responses to light I. Involvement of GABA[A], benzodiazepine and GABA[B] receptors. *J. Neurosci*. 9: 2858-2865.

23.   Rea, M.A.; Glass, J.D.; Colwell, C.S. (1994) Serotonin modulates photic responses in the hamster suprachiasmatic nuclei. J. Neurosci. 14: 3635-3642.

24.   Rusak, B; Zucker, I. (1979) Neural regulation of circadian rhythms. *Physiol. Rev.* 59: 449-526.

25.   Sawaki, Y.; Nihonmatsu, I.; Kawamura, H. (1984) Transplantation of the neonatal suprachiasmatic nuclei into rats with complete bilateral suprachiasmatic lesions. *Neurosci. Res.* 1: 67-72.

26.   Swanson, L. W.; Cowan, W. M.; Jones,E. G. (1974) An autoradiographic study of the efferent connections of the ventral geniculate nucleus in the albino rat and the cat. *J. Comp. Neurol.* 156: 143-164.

27.   Shibata, S.; Liou, S. Y.; Ueki, S. (1986) Influence of excitatory amino acid receptor antagonists and of baclofen on synaptic transmission in the optic nerve to the suprachiasmatic nucleus in slices of rat hypothalamus. *Neuropharmacology* 28: 403-409.

28.   Takahashi, J. S.; DeCoursey, P. J.; Bauman, L.; Menaker. (1983) Spectral sensitivity of a novel photoreceptive system mediating entrainment of mammalian circadian rhythms. *Nature* 308: 186-188.

29.   van den Pol, A. N.; Tsujimoto, K. L. (1985) Neurotransmitters of the hypothalamic suprachiasmatic nucleus: immunocytochemical analysis of 25 neuronal antigens. *Neuroscience* 15: 1049-1086.

Microparticle Bioluminescence

Robert L. Armstrong
Professor
Department of Physics

New Mexico State University
P O Box 3D
Las Cruces, NM 88003

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

August, 1996

Microparticle Bioluminescence

Robert L. Armstrong
Professor
Department of Physics
New Mexico State University

Abstract

The summer research period was devoted to several tasks: an experimental study of the operation and stability of an electrodynamic balance of potential use in the detection of bioluminescent emission from microparticles; a number of discussions on research activities with colleagues; participation in a research conference held each summer at the Edgewood laboratory; and, establishing contacts with AFOSR with regard to ongoing research interests.

Microparticle Bioluminescence

Robert L. Armstrong

Introduction

From approximately mid-may through mid-July, 1996, I collaborated in a research project with Dr. Burt Bronk at the Armstrong Laboratory facility in Edgewood, MD. There is significant current interest in the detection of hazardous chemical and biological material. Dr. Bronk was investigating the phenomenon of bioluminescence as an example of problems encountered in the detection of biological substances. In a research effort conducted last summer, Dr. Bronk and I successfully measured bioluminescence in micromolar samples of material; however, the sensitivity of our detection system was too low to obtain measurable emission from micron-sized droplets of the material (which contain picomolar quantities of the biological material). A number of possible improvements suggested themselves during last summer's work, including, most importantly, the design of a small f-number optical detection system. During this summer's effort, it was decided to concentrate on this system.

Electrodynamic Trap

In order to perform spectroscopic measurements on a micron-sized sample, consideration must be given to the stable suspension of the sample. A number of schemes have been successfully employed to accomplish this, including the so-called electromagnetic (1) and the electrodynamic (2) traps. The former trap uses direct photon pressure from a vertically-pointing laser beam to balance the gravity force; further, if the laser source is operating in the conventional TEM_{00} mode, lateral stability is also achieved and the particle is confined within a spatial region comparable in size to the focal volume of the trapping laser. However, the electromagnetic trap requires that the trapped particle be highly transparent at the laser wavelength; if this is not the case, heating (photophorectic) effects overpower photon pressure, the particle becomes dynamically unstable, and is ejected from the trap.

Alternatively, the electrodynamic trap does not require that the trapped particle be transparent but relies on electric forces acting on the particle to achieve trapping. In order to accomplish this, the particle is first charged and then injected into a spatial region containing a combination of electrodes providing ac and dc electric fields. Many electromagnetic trap designs employ solid electrode caps together with cylindrical electrodes in the equatorial plane of the particle. Hence, if the particle emits luminescence, the light may be collected only from a small fraction of the total solid angle; in other words, the collection optics

possesses a large f-number and is unsuitable for the detection of weak bio- luminescent emissions expected from microparticles.

Some electrodynamic trap designs involve a more open electrode structure, and are more readily incorporated into low f-number optical detection systems. In particular, a trap has been successfully tested that consists of two wire loops upon which are impressed a superposition of ac and dc voltages (3). The stability of this trap is a function not only of the magnitude of the impressed voltages but also of the diameter of the loops and their separation. Since the openness of this trap recommended itself to our need for light collection from a large solid angle, we elected to build and test such a loop trap during this past summer.

A loop trap was constructed and subjected to a large number of tests during the summer research period. Difficulties were experienced with both the trap stability and with the microparticle injection system (consisting of a piezoelectrically pulsed spray nozzle and a charging electrode). Toward the end of the research period, these difficulties were overcome and test glycerol/water mixture microparticles were stably levitated in the trap.

The next phase of this project is the design and construction of an environmental chamber. Such a chamber is necessary since the bioluminescent particles typically consist of an aqueous solution of biological material that evaporates quickly under ambient room conditions. What is needed is an enclosed chamber (containing the trap) within which the microparticle may be stably levitated over long periods (tens of minutes) during which the temporally decaying bioluminescent emission is measured. At the termination of the research period, a candidate design for such a chamber was discussed. Fabrication and testing of this chamber will occur later.

Other Activities

In addition to working with Dr. Bronk on the bioluminescent experiment discussed above, I also had discussions with several colleagues at the Edgewood laboratory regarding their ongoing research activities. These discussions, which centered on problem areas in their research, were (I hope) helpful, since I was able to provide a trained ear. I also participated in a conference on aerosols, held each summer at Edgewood, contributing a research paper describing my recent research activities in this area. Finally, I was able to make contact with individuals in AFOSR having an interest in the area of biological detection; these contacts provided Dr. Bronk and I with valuable information important to future biodetection research efforts.

References

1. A. Ashkin and J. D. Dziedzic, Appl. Opt. 20, 1803 (1981).

2. S. Arnold and L. M. Folan, Rev. Sci. Instrum. 58, 1732 (1987).

3. E. J. Davis et al, Rev; Sci. Instrum. 61, 1281 (1990).

LACK OF EFFECT OF ULTRAWIDEBAND RADIATION ON PENTYLENETETRAZOL-INDUCED CONVULSIONS IN RATS

Maureen E. Bronson
Associate Professor
Department of Pharmaceutical Sciences

Wilkes University
School of Pharmacy
Wilkes-Barre, PA 18766

LACK OF EFFECT OF ULTRAWIDEBAND RADIATION ON PENTYLENETETRAZOL-INDUCED CONVULSIONS IN RATS

Maureen E. Bronson
Associate Professor
Department of Pharmaceutical Sciences
Wilkes University School of Pharmacy

Abstract

The purpose of the present study was to determine whether exposure to high peak power, short pulses of ultrawideband (UWB) electromagnetic radiation would affect susceptibility to convulsions caused by pentylenetetrazol (PTZ). The ED_{99} for PTZ-induced convulsions was determined to be 89.174 mg/kg, and this was the dose administered to 2 groups of rats, one sham exposed and one exposed to UWB produced by a Kentech pulse generator. Parameters for the Kentech were 55kV/m peak E field, 30 ps rise time, 1.8 ns pulse width and 1 kHz pulse repetition rate. Two additional groups were administered saline injections and received either a sham or UWB exposure. UWB had no visible effect on saline injected animals, nor did it affect latency to convulsion, length of convulsion or lethality produced by PTZ when these rats were compared to sham exposed animals that received PTZ.

5-2

LACK OF EFFECT OF ULTRAWIDEBAND RADIATION ON PENTYLENETETRAZOL-INDUCED CONVULSIONS IN RATS

Maureen E. Bronson

Introduction

Ultrawideband (UWB) radar systems are being used increasingly in the military for many applications, including locating buried explosive mines {1}. Because the U.S. Air Force, Army and Navy are all developing and testing UWB sources {2, 3, 4, 5}, it is imperative that research be conducted on possible biological effects of these systems. Although there has been at least one report theorizing that UWB should be biologically active {6}, all of the studies conducted with UWB to date have found no biological effects of this particular type of nonionizing radiation. For example, rats exposed to high-peak-power electromagnetic pulses generated by a Hindenberg 2 UWB transmitter showed no changes in blood chemistry, c-fos expression in brain, swimming performance or a functional observational battery composed of behavioral tests designed to serve as a screen for environmental and chemical toxins {7}. Primates exposed to the same UWB parameters also failed to show behavioral effects when tested in the

Primate Equilibrium Performance task {8}. Recently, Jauchem, Frei, Ryan, Merritt and Murphy {9} have extended these studies using a different UWB source (Bournlea 3153 pulse generator) and have found that exposure to fast-rise-time UWB electromagnetic pulses had no effect on heart rate or blood pressure in rats. The possibility that UWB could alter biological membranes has been suggested {6}, and if this is the case, then one might expect that UWB would increase seizure activity in susceptible individuals, because seizure activity is essentially a depolarizing shift characteristic of a large excitatory postsynaptic potential. Military personnel face the possibility of being exposed to chemical warfare agents in the field, some of which are known to cause convulsions. If the same personnel are also exposed to UWB, then it is possible that they would be more susceptible to convulsions caused by chemicals such as Soman. PTZ was chosen as the chemical convulsant for this experiment because it has a rapid onset and short duration of seizure activity. It is also a pure convulsant, lacking the severe cholinergic effects of drugs such as Soman.

Methodology

Eighty male Sprague-Dawley rats (Charles River Laboratories, Portage, MI) weighing between 380-460 g. were used in this study. The animals involved in

this study were procured, maintained and used in accordance with the Animal Welfare Act and the "Guide for the Care and Use of Laboratory Animals" prepared by the Institute of Laboratory Animal Resources - National Research Council. Formal approval, in advance, was received from the Armstrong Laboratory Animal Use Committee to conduct these experiments. Animals were housed in polycarbonate cages with free access to food and water, and maintained on a 12 h/12 h, light/dark cycle (lights on at 0600) in a climate-controlled environment (ambient temperature $24 \pm 1^\circ$ C).

A dose-effect curve was first conducted for PTZ (Sigma Chemicals, St. Louis, MO) to determine the ED_{99} for convulsions. Thirty rats (6 rats/dose) received i.p. injections of PTZ that had been serially diluted to the following doses: 60, 69.29, 80.03, 92.42 and 106.76 mg/kg. The PTZ was dissolved in saline and the injection volume was 1 ml/kg. Fresh solutions of PTZ were prepared daily and the filled syringes were kept on ice. Time to convulsion and lethality were recorded.

The day following determination of the ED_{99}, 40 rats were divided into 4 groups of 10 rats each and were treated as follows. One group received the ED_{99}

convulsant dose of PTZ 20 sec before a 2 min exposure to UWB produced by a Kentech pulse generator and a second group received an injection of saline before being exposed to the Kentech. These groups are referred to as the exposure groups. A third group received the ED_{99} convulsant dose of PTZ with no exposure to UWB and a fourth group received saline with no exposure to UWB. These groups constituted the sham exposure group because they were placed in the chamber and subjected to a pulsed tone that normally accompanies UWB exposure, but the UWB source was not turned on. A random number generator was used to determine group assignments. Parameters for the Kentech were as follows: 55 kV/m peak E field, 30 ps rise time, 1.8 ns pulse width and 1 kHz pulse repetition rate. A video camera was used to record all rats throughout the injection, exposure and post-exposure period. Latency to first response (typically a brief head or whole body jerk), latency to tonic-clonic convulsion, duration of each convulsion, number of convulsions and number of deaths were recorded.

Data Analysis

Probit analysis was used to determine the ED_{99} convulsant dose of PTZ. T-tests were used to analyze convulsion latencies and durations. Number of convulsions and number of deaths were analyzed using a two-factor General Linear Models

5-6

(GLM) Analysis. Significance was set at $p < 0.05$.

Results

PTZ ED_{99}.   The ED_{99} for PTZ was 89.174 mg/kg (95% CI = 79.026 < ED_{99} < 162.235). At the lowest dose of 60 mg/kg, only 1/6 rats convulsed, and the convulsion was characterized by intermittent jerking movements rather than a full blown tonic-clonic convulsion. No rats died at the 60 mg/kg dose. The dose of 69.29 mg/kg also produced intermittent jerking movements rather than full tonic-clonic convulsions, but 1 rat died at that dose. At 80.03 mg/kg 3 rats died, despite the fact that convulsions were not severe. Full-blown, continuous tonic-clonic convulsions were seen at 92.42 and 106.74 mg/kg and all but 1 rat died after receiving those doses. Interestingly, the rat that lived was in the highest dose group.

Exposure. All rats that received the ED_{99} dose of PTZ convulsed. As shown in Fig. 1, there was no difference between latency to first seizure response for the sham and UWB groups that received PTZ (mean+SEM = 59.8±3.2 sec for sham and 58.6±3.9 sec for UWB rats). One of the sham rats was inadvertently administered an intravenous injection, as evidenced by the fact that he started convulsing as soon as the needle was removed. Data from this rat are therefore

not included. There was also no difference in latency to full convulsion (Fig. 2; mean±SEM = 65.3±2.5 sec for sham and 64.7±3.8 sec for UWB rats) or end of seizure activity/death (Fig. 3; mean±SEM = 883.0 ±201.9 sec for sham and 1025.7±166.3 sec for UWB rats). All but 3 of the PTZ-treated rats (1 sham and 2 UWB) died within an hour. Time to death was highly variable, ranging from less than 5 minutes to approximately 34 minutes. Fig. 4 shows no significant differences for duration of seizure (mean±SEM = 798.3±201.8 sec for sham and 912.6±165.6 sec for UWB rats). As can be seen in Fig. 5, none of the saline treated animals convulsed and there was no significant difference between number of convulsions in the PTZ-treated rats (mean±SEM = 1.33±0.24 convulsions for sham and 1.9±0.41 convulsions for UWB rats), as evidenced by the fact that there was an effect of drug (p=0.0001) but not exposure (p=0.2523).

Discussion

The ED_{99} for PTZ convulsions as determined in this study is very close to the reported LD_{50} for PTZ (100 mg/kg) after s.c. injections in the same strain of rat {10}. In that study, all rats had intense convulsions at 100 mg/kg. In the current study, all of the PTZ rats convulsed at 89.17 mg/kg and 84% of the PTZ-injected rats died, regardless of whether they received UWB or sham exposure. Thus, it

5-8

would appear that PTZ is more potent after an i.p. versus a s.c. injection, as evidenced by the fact that 84% of the rats in this study died at an i.p. dose of 89.17 mg/kg whereas only 50% died at a s.c. dose of 100 mg/kg {10}.

UWB had no apparent effects on saline-treated rats. This finding extends those of other investigators who have found no adverse effects of UWB on biological systems {7, 8, 9}. UWB exposure also did not alter the effects of PTZ, as may have been expected if UWB causes membrane perturbations, as suggested by Albanese et al. {6}. Latency to convulsion was the same in both groups and correlates well with latency data reported by Goldman, Berman, Hazlett and Murphy after i.p. injection of PTZ {11}. In Goldman et al. {11}., however, an ED_{50} of 42 mg/kg was reported, whereas in the present study, only 17% of the rats convulsed at the lowest dose of 60 mg/kg during determination of the ED_{99}. The animals in the referenced study were also male Sprague-Dawley rats purchased from the same source and weighing approximately the same as our rats (i.e., 380-430 g). In Goldman et al. {11}, however, convulsions were detected by EEG, whereas in the current study convulsions were detected by visual monitoring. It is possible, therefore, that seizure activity was taking place at the lower PTZ doses, but that we were unable to detect it.

5-9

In summary, rats exhibited typical seizure activity after being treated with PTZ, whether exposed to UWB or sham environment. UWB had no apparent effects on saline treated rats.

Acknowledgments.   The author would like to thank all of the Armstrong Laboratory personnel who were involved in this project: Dr. Michael Murphy, Ms. Stephanie Miller and SSgt Max Williams.

References

1. D. Herskovitz. Wide, wider, widest. J. Electronic Defense, July, 1995, 51-58.

2. F.J. Agee, D.W. Scholfield, W. Prather and J.W. Burger. Powerful ultra-wide band RF emitters: status and challenges. Proc. SPIE - Int. Soc. Optical Eng. 2557:98-19, 1995.

3. R.A. Koslover, D.E. Voss, C.B. Jerome, R.J. Guarnieri, D.A. Slemp and J. D. Karasz. A design study of high-gain, high power, fast-risetime, short-pulse, ultra-wideband (UWB) antennas. USAF Phillips Laboratory, Kirtland AFB, NM, Tech. Rep. PL-TN-93-1021, August, 1993.

4. D. Birx and J.R. Fishback. High power, 10 kHZ repetition rate ultra-wideband source development at the US Army's Missile Command.   IEEE MTT-S International Microwave Symposium Digest 3:1605-1608, 1992.

5. A.T. Richardson and E.N. Clouston. Broadband aerodynamic antennas for the advanced radar missile scorer. In IEEE Colloquium on Antenna and Propagation Problems of Ultrawideband Radar, January 12, 1993, London, UK, pp. 4/1-5.

6. R. Albanese, J. Blaschak, R. Medina and J. Penn. Ultrashort electromagnetic signals: biophysical questions, safety issues and medical opportunities. Aviat. Space Environ. Med. 65 (suppl 5):A116-A120, 1994.

7. T.J. Walters, P.A. Mason, C.J. Sherry, C. Steffen and J.H. Merritt. No detectable bioeffects following acute exposure to high peak power ultrawide band electromagnetic radiation in rats. Aviat. Space Environ. Med. 66:562-567, 1994.

8. C.J. Sherry, D.W. Blick, T.J. Walters, G.C. Brown and M.R. Murphy. Lack of behavioral effects in non-human primates after exposure to ultrawideband electromagnetic radiation in the microwave frequency range. Radiat. Res. 143:93-97, 1995.

9. J.R. Jauchem, M.R. Frei, K.L. Ryan and M.R. Murphy. High-peak-power pulses from TEMPO and ultrawideband sources: heart rate and blood pressure responses in rats. Proceedings of the Bioelectromagnetics Society, Seventeenth Annual Meeting, p. 151, 1995.

10. T. Kadar, G. Cohen, R. Sahar, D.Alkalai and S. Shapira. Long-term study of brain lesions following soman, in comparison to DFP and metrazol poisoning. Human Exp. Toxicol. 11:517-523, 1992.

11. H. Goldman, R.F. Berman, J. Hazlett and S. Murphy. Cerebrovascular responses to pentylenetetrazol: time and dose dependent effects. Epilepsy Res. 12:227-242, 1992.

FIGURE LEGENDS

Fig. 1. Latency to first evidence of seizure activity in UWB versus sham rats that received 89.174 mg/kg PTZ. Values represent the mean number of seconds $\pm$ SEM.

Fig. 2. Latency to first full convulsion in UWB versus sham rats that received 89.174 mg/kg PTZ. Values represent the mean number of seconds $\pm$ SEM.

Fig. 3. Time from onset of seizure activity until end of seizure activity or death, whichever occurred first, in UWB versus sham rats that received 89.174 mg/kg PTZ. Values represent the mean number of seconds $\pm$ SEM.

Fig. 4. Duration of seizure activity in UWB versus sham rats that received 89.174 mg/kg PTZ. Values represent the mean number of seconds $\pm$ SEM.

Fig. 5. Number of convulsions in UWB versus sham rats that received 89.174 mg/kg PTZ. Values represent the mean number of convulsions $\pm$ SEM.

Fig. 1
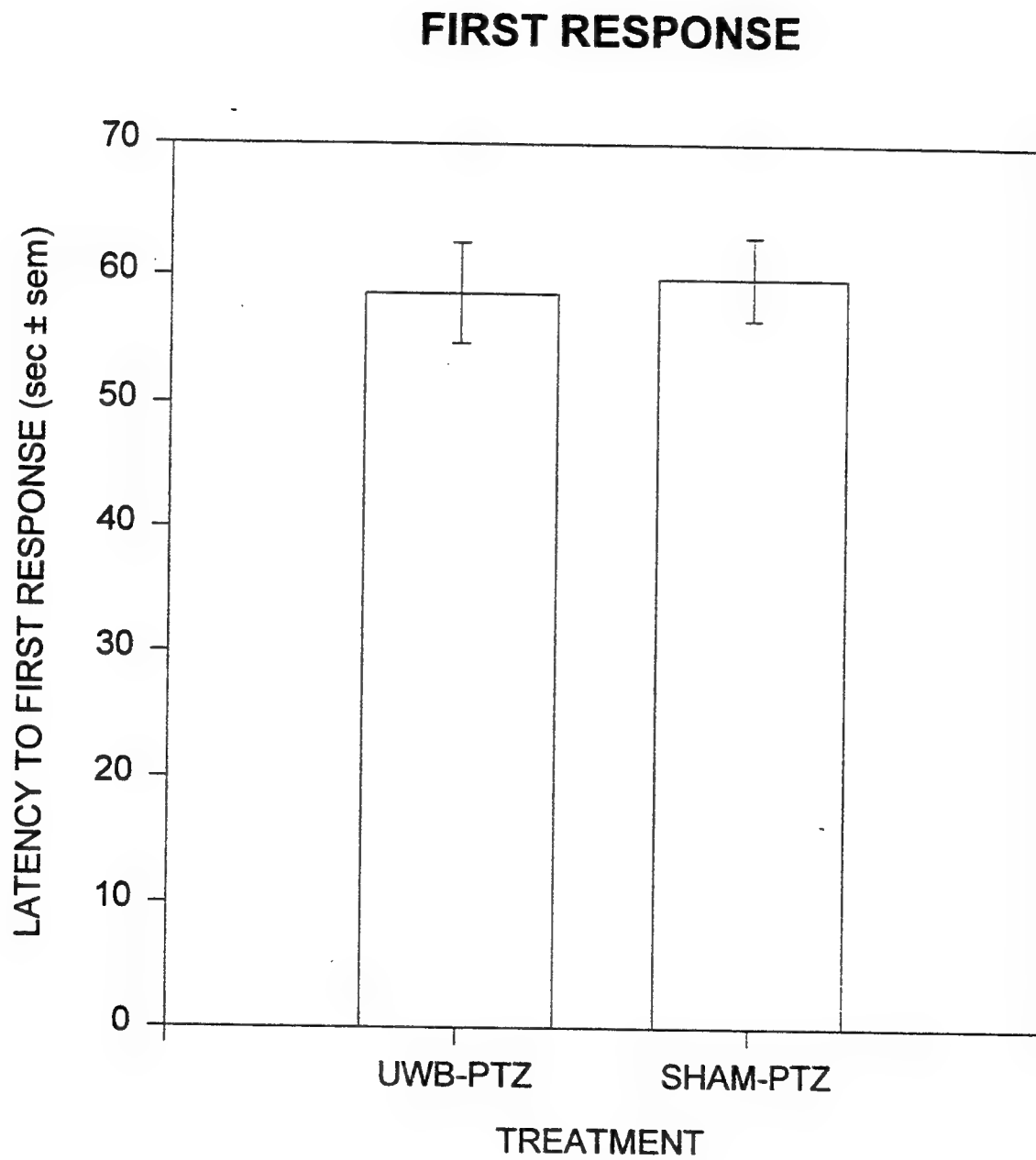
FIRST RESPONSE
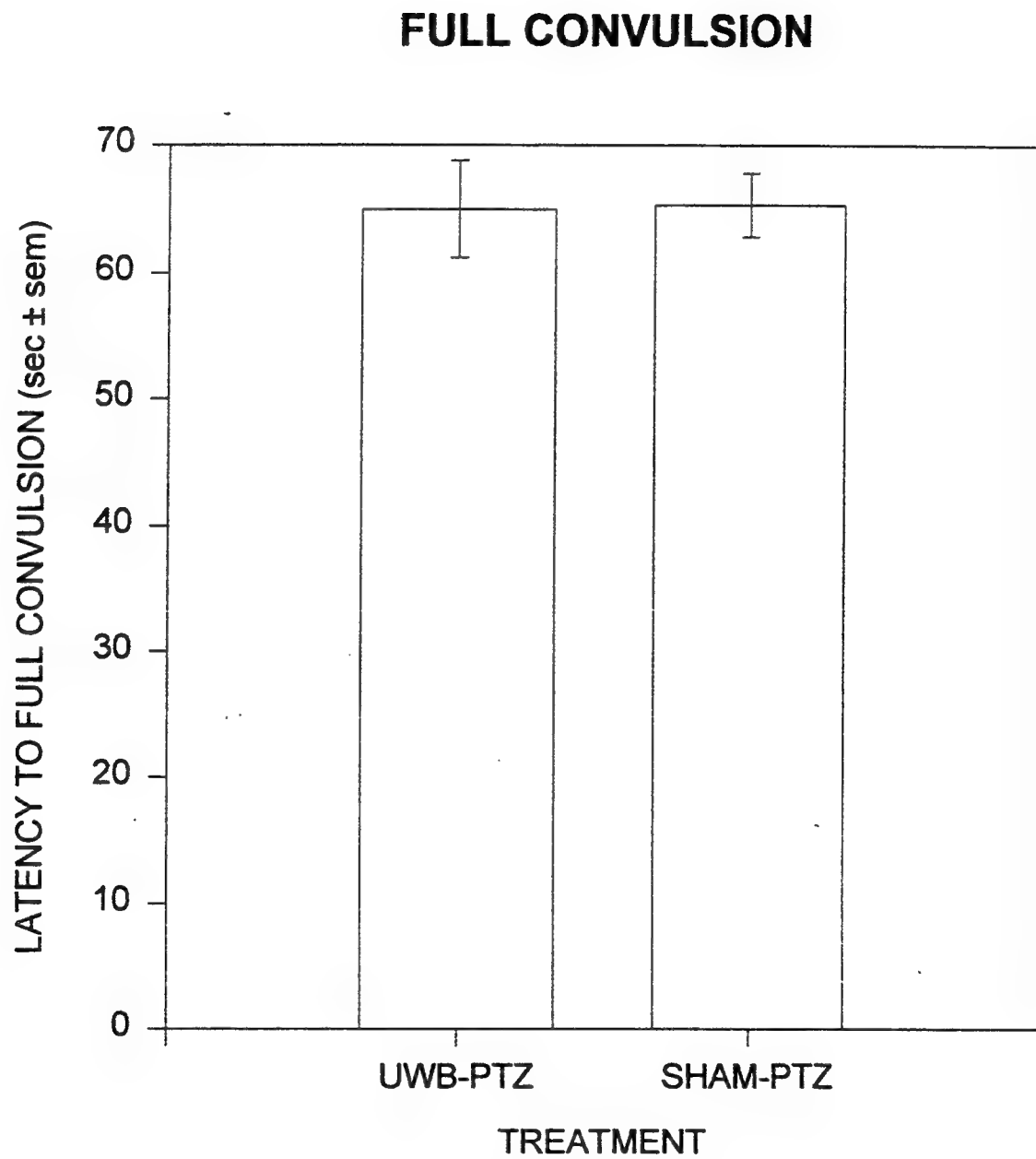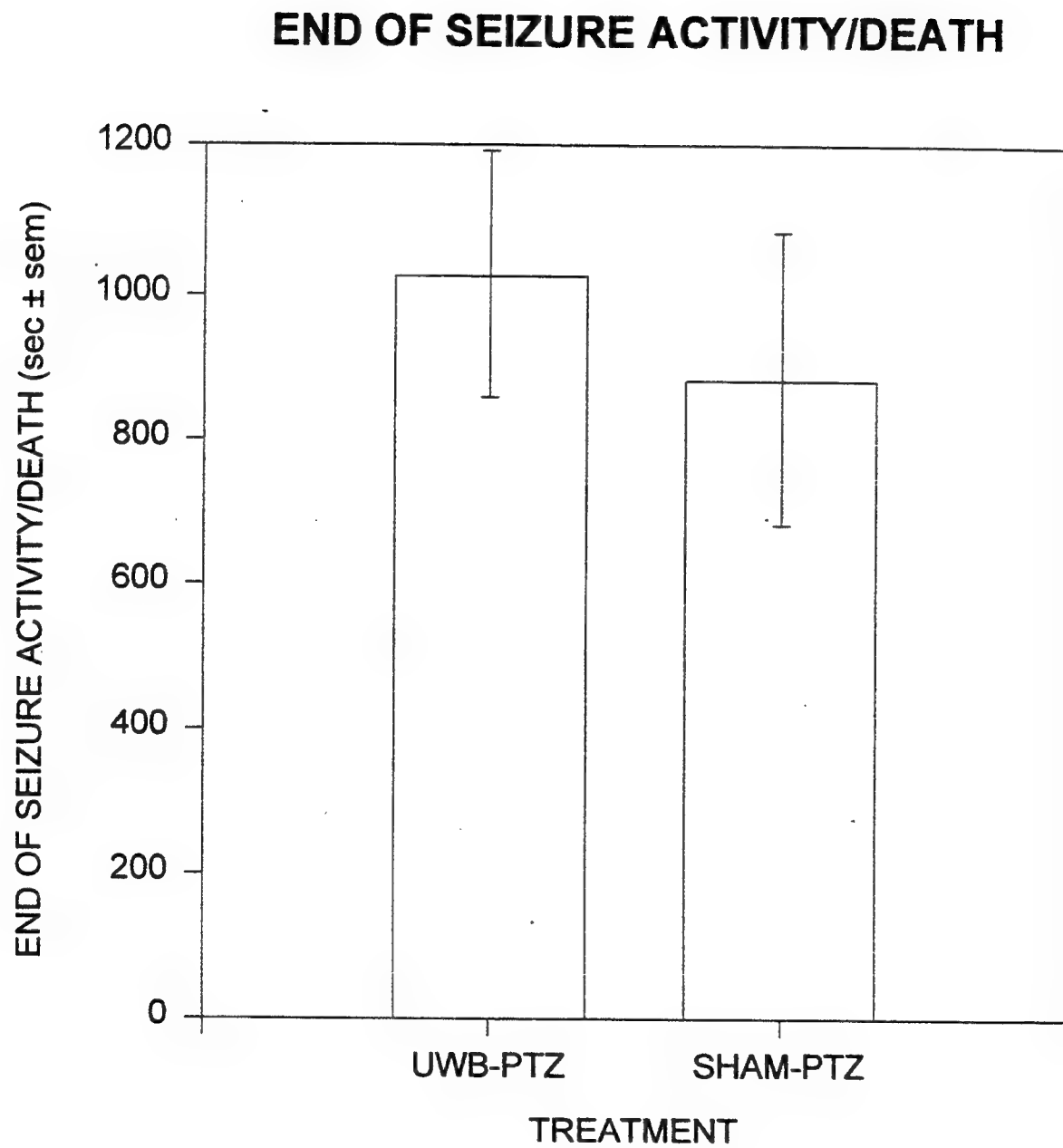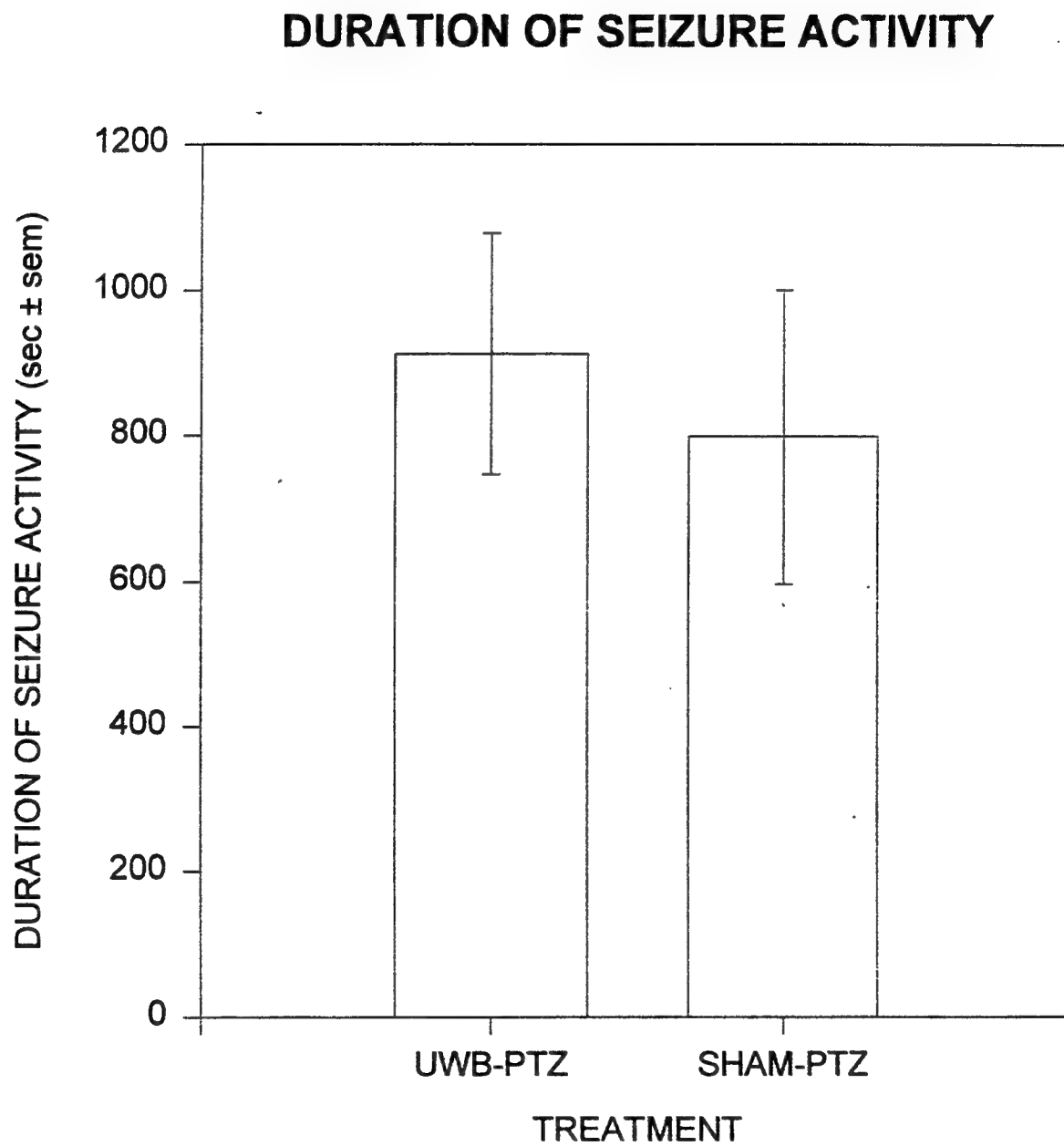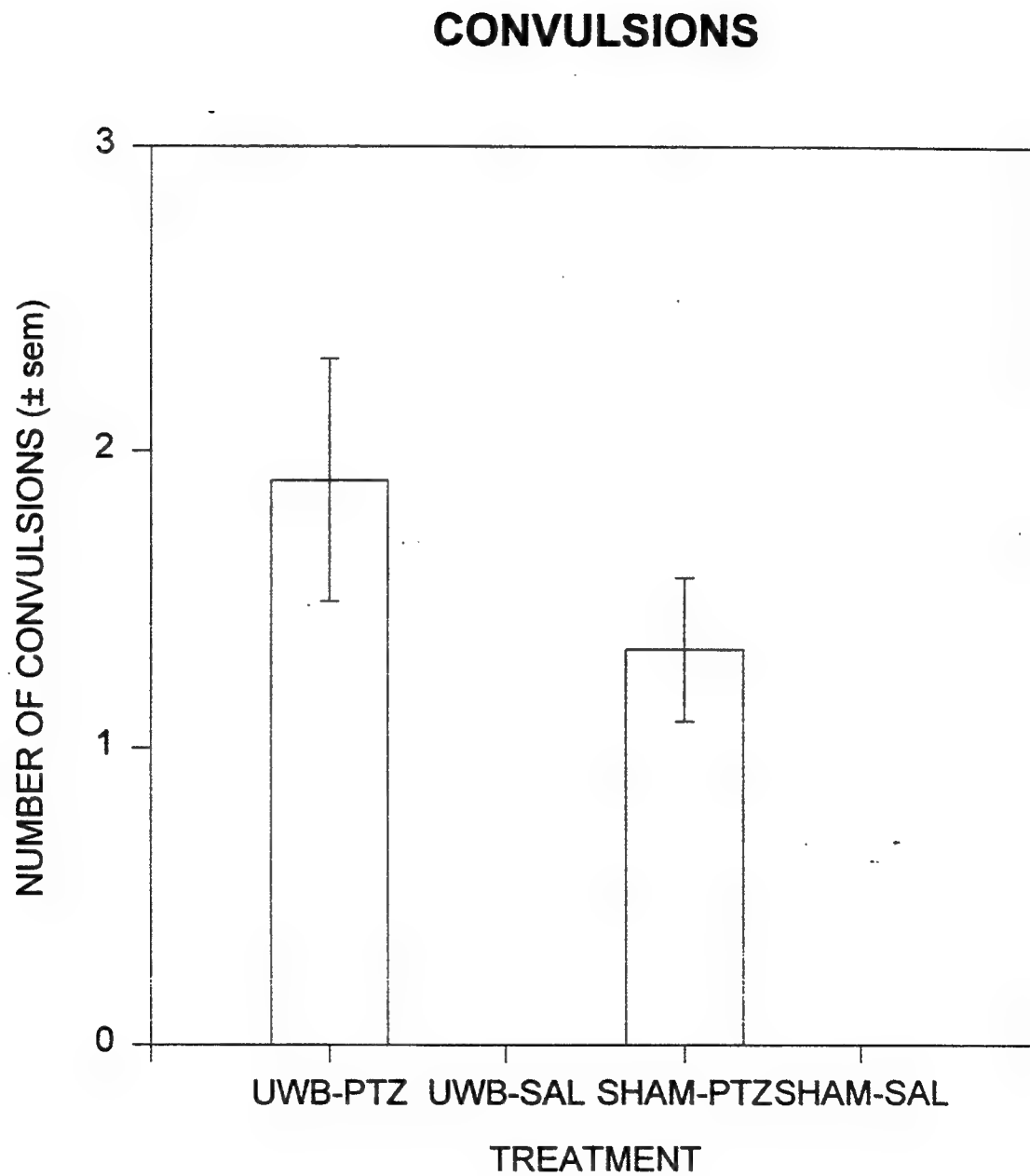
Fig. 2

FULL CONVULSION

Fig. 3

END OF SEIZURE ACTIVITY/DEATH

Fig. 4

DURATION OF SEIZURE ACTIVITY

Fig. 5

CONVULSIONS

Assessment of the Reliability of Ground-Based Observers for the Detection of Aircraft

Marc L. Carter
Assistant Professor
Department of Psychology

University of South Florida
4202 East Fowler Avenue, BEH 339
Tampa, Florida   33620-8200

and

Jason McCarley
Department of Psychology
University of Louisville
2301 South Third Street
Louisville, KY   40208-1802

Assessment of the Reliability of Ground-Based Observers for the Detection of Aircraft

Marc L. Carter
Assistant Professor
Department of Psychology
University of South Florida

Jason McCarley

Department of Psychology
University of Louisville

Abstract

In situations in which ground-based lasers are propagated through the atmosphere, either for entertainment or scientific pursuits, there is the chance that aircrew may be exposed to the beam. In most cases this exposure would not be eye-hazardous, but the effects of flashblindness and veiling glare can nonetheless impair mission performance, with potentially catastrophic consequences. In most situations where such lasers are employed, ground-based observers attempt to identify aircraft that are in or near the beam path; occasionally these observers are aided by FAA radar feeds that can assist them in locating these aircraft. In this study we attempt to determine the effectiveness of observers in the detection of aircraft under a variety of conditions, including day versus night, and with and without the assistance of a radar feed. Preliminary data collected at Sandia National Labs in Albuquerque, NM, suggest several points. First, detection range is very much greater at night than in the day, probably due to the high contrast between the aircraft and night sky from aircraft lighting, and the increased visual sensitivity of the observers in scotopic viewing. Second, the assistance of a radar feed for daytime observation is important in aircraft detection, not so much to increase the range at which the aircraft is visually acquired, but to increase the likelihood that the aircraft will be detected at all. In further analysis of the complete data set we will examine the impact of various ground and sky conditions that can mitigate the performance of the observers.

Assessment of the Reliability of Ground-Based Observers for the Detection of Aircraft

Marc Carter and Jason McCarley

As lasers have come into widespread military application, enormous effort has been dedicated to ensuring their safe use. Much of this has aimed at assessing the biological effects of ocular exposure to laser, and at determining eye-safe levels of exposure for a variety of lasers, both battlefield and other. It is well established, though, that levels of exposure too weak to cause biological damage can impair visual function (Thomas, 1994). Either glare--scattered light veiling a scene (Pulling, Wolf, Sturgis, Vallancourt, & Dolliver, 1980; Vos, 1984)--or flashblindness--the loss of sensitivity which follows exposure to intense light (Brown, 1965)--can cripple visual performance. Both effects are temporary, but consequential. For pilots, momentary functional blindness can devastate, most importantly at mission-critical junctures, when the aircraft is more likely to be near the ground (targeting, landing and takeoff).

For some time glare and flashblindness have been of great interest to the military, increasingly as the potential for laser eye exposure has increased due to the potential for lasers to be wielded as weapons by enemies or the actual use of lasers as tools by friends. More recently, however, laser exposure to civilian aircrew has become a concern as the number of lasers employed in entertainment, advertising, and scientific applications has increased (Aviation Week & Space Tech, Sept. 26, 1994; Weiss, 1996). The potential for danger became a reality when a commercial airline pilot, upon departure from Las Vegas' McCarran International Airport, was temporarily blinded by a casino laser, and forced to turn control of his aircraft over to the copilot (Scott, 1995). This event, having been preceded by complaints of approximately fifty similar

incidents, prompted the FDA to temporarily prohibit laser light shows in the city of Las Vegas, and to consider a similar, nationwide moratorium (Air Line Pilot, 1996). But such a prohibition could be fully effective only if it also banned use of atmospherically transmitted scientific lasers, whose high-intensity, low-divergence beams remain hazardous over greater distances than the typically low-power beams employed by commercial displays (Weiss, 1996). A summary proscription of outdoor lasers, then, would confiscate a tool of both science and commerce. This is a cost to be incurred if necessary, but to be avoided otherwise.

One popular safeguard against the dangers of outdoor lasers is the simple use of ground-based observers to monitor the airspace through which the lasers are deployed, and to control the firing of lasers when there are aircraft in or approaching the beam path. A more sophisticated and potentially more reliable operation employs a radar feed from (typically) FAA radar located at an airport in the vicinity. Such a system, although perhaps better than simple observation, still suffers from several defects, among them the inability of FAA radar to accurately locate aircraft near the ground (where they most obviously are during takeoff and landing, perhaps the most critical portions of any flight) as well as in the "blind spot" inherent in ground-based radar systems (Weiss, 1996), and the temporal lag involved in registration of an aircraft's location on the radar display. The combination of a radar feed along with visual observation offers the best solution present to date, in that a radar operator can direct the observers to a region of airspace in which an aircraft is indicated by the radar; the visual observers then are able to very accurately determine the current location of the aircraft with respect to the beam path.

Under most current circumstances, however, unassisted observers are likely to be the sole safeguard against laser exposure to commercial and private aircraft, in spite of the fact that their

reliability has not been assessed. Past research has generally examined the ability of observers to acquire aerial targets appearing against the relatively uniform backdrop of a daylight sky (e.g., Akerman & Kinzly, 1979; O'Neal, Armstrong, & Miller, 1988), and models of visual target acquisition have likewise, explicitly or implicitly, assumed photopic illumination (e.g., Akerman & Kinzly, 1979; Koopman, 1986). For several reasons, the results of these efforts cannot be easily generalized to many of the conditions under which outdoor lasers are deployed. Under photopic illumination, a distant aircraft at threshold contrast will likely appear as a still, dark patch, undistinguished by color or motion (Akerman & Kinzly, 1979; Koopman, 1986). On the other hand, color, motion, and flicker might be the most salient source of information about the approach of a night-flying aircraft bearing an array of identification, landing, and stroboscopic anticollision lights. This information, along with increased visual sensitivity that accompanies a decline in ambient illuminance, may well allow acquisition ranges measured at night to exceed those obtained in daylight. Unfortunately, the dark adapted eye is most susceptible to glare (Sturgis & Osgood, 1982) and nighttime performance, vulnerable to the ambient light of a city or airport, may be volatile.

Daytime and nighttime target acquisition are likely to represent fundamentally different visual abilities, such that performance between tasks will differ substantially. The present research will therefore examine the ability of visually unaided ground-based observers to monitor airspace under photopic and scotopic conditions, and with consideration for the terrain over which search is conducted. Additionally, acquisition ranges will be compared between conditions in which observers work in concert with a radar operator, and those in which observers freely scan an assigned area of space. Results will be used to assess the efficacy of such observers as a safeguard

against exposure of aircrew to atmospherically propagated lasers.

Method

Observers

Eight male volunteers were recruited from among the employees of Sandia National Laboratory, and paid for their participation. The observers were screened to ensure normal color vision, and visual acuity of 20/20 or better, and ranged in age from 33 to 35 years. The observers used in this study had a variety of experience at spotting for aircraft, some having never performed the task prior to testing, and others having 30 or more hours experience.

Procedure

Data were (and continue to be) collected at Sandia National Laboratories' LAZAP Facility, located on the Sandia Military Reservation, Albuquerque, New Mexico. The facility is the site of a high-powered, low beam-divergence, ruby laser, used for the calibration of on-orbit DOD and DOE satellite payloads. The data are being collected apart from the Facility's normal operations, and do not involve any use of lasers.

Subjects will be stationed approximately 3 m above ground level, on a rooftop near the Laboratory's outdoor laser, during times at which the laser is not operational. For day observation sessions, subjects are encouraged to use appropriate sunscreen. Test sessions last approximately four hours, and are divided into one-hour shifts of data collection separated by periods of rest, with a long rest (approximately one-half hour) after two hours.

A simple 2 (unaided vs. aided search) by 2 (photopic vs. scotopic illumination) factorial

design will be used to analyze data. In all conditions, observers are asked to indicate when they visually acquire the target. In the free scanning (unassisted) conditions, the airspace surrounding the point of observation is divided into hemifields facing either north and south or east and west. Throughout a single shift, observers monitor an assigned hemifield of airspace, and report any aircraft they acquire within that airspace. In aided scanning conditions, observers monitor 360 degrees of airspace each shift, but are given the range, azimuth, and altitude, of targets entering the monitored airspace. Daylight shifts are run between 10:00 AM and 2:00 PM, in order to minimize cues from glint off the aircraft or having the aircraft in line with the position of the sun.

In the assisted condition, aircraft location information is provided by a feed from the FAA radar at the Albuquerque airport, and are noted by a Radar Airspace Monitoring System (RAMS) operator. Data collected include the time of day at which each target appears, cloud cover and visibility at that time, the range, azimuth, and elevation of the target when notification is made, and then the range, azimuth and elevation of the target when it is detected. If a target goes undetected, its nearest approach to the observer is recorded. The straight-line distance from the observer at which a target is acquired is the most immediate measure of performance.

Preliminary Results, and Some Discussion

At the writing of this report, data collection was not complete. However, data from approximately half of the observers is available. For this preliminary analysis, we use "clean" data; that is, aircraft detection that relied on anything other than simple detection was eliminated from the summary. Hence, no detections that were aided by glint, contrails, or sound, or detections

that were hampered by clouds, veiling glare from nearby lights, or obscured by buildings, are included. We also eliminated data from aircraft that were near take-off or landing, although in the night conditions detection via landing lights from distant aircraft is allowed. The data are also presented as linear miles from the observer, without consideration of azimuth or elevation; those parameters will be considered in the complete analysis. The data are classified as "hits," which is the point at which the observer reported spotting the aircraft, or "misses," in which case the data reflect the aircraft's closest point of approach to the observer.

First, the data from the day conditions:

	Assisted			Unassisted		
	Mean	SEM	N	Mean	SEM	N
Hits	11.37	.681	63	8.03	.140	28
Misses	13.07	.630	16	12.11	.313	21

There are a couple of points of relevance in these data. First, note that observers were detecting aircraft at greater distances in the assisted rather than unassisted conditions. This clearly demonstrates that for daylight observations the assistance of the radar feed is to be preferred. Another aspect of these data is the difference in the quantity of detections between the two conditions, but no firm conclusions can be drawn from this since, although we attempted to equate the amount of time the observers spent in each condition, we obviously had no control over the air traffic during those sessions. It is possible, however, to remark on the difference in *proportion* of aircraft detected versus missed. Of the 79 aircraft present during the assisted sessions, the observers spotted very nearly 80%, whereas in the unassisted conditions they only

spotted 57%. This also lends support for the benefit of having assistance in searching for aircraft.

Next, the data from the night observations:

	Assisted			Unassisted		
	Mean	SEM	N	Mean	SEM	N
Hits	23.49	.147	49	24.60	.281	38
Misses	22.72	2.557	3	21.00	1.38	6

The first thing to note about these data is that there is only a small (but nonetheless reliable) difference in mean detection range for the two conditions (assisted versus unassisted), which at first may seem surprising. However, this is almost surely due to the much greater ability of the observers to detect aircraft by means of lights (either anticollision or landing lights) at night, when the observers sensitivity is greater and the aircraft's contrast against the sky is high, regardless of whether or not the observer is directed to the area of sky in which the aircraft can be found. They are simply much easier to see at night, and the RAMS feed is of much less import. Also note that the proportions of hits versus misses is much the same for the two conditions, although there is a slightly higher proportion in the unassisted condition. The last, and perhaps most salient, point is to compare the data from the day and night sessions. It is clear that the effective range of the observer is very much greater (at least twice as great) at night than in the daytime. Although not surprising, it is important to note: observers simply see the aircraft farther away at night, due to the presence of landing lights or anticollision strobes. When the contrast between the aircraft and the background is lower, such as in the daylight conditions, performance suffers.

To sum up, we have learned from the preliminary examination of the data that the RAMS operator is very important during the daylight, both to increase the range at which aircraft may be detected, and to increase the likelihood that an aircraft will be detected. At night, however, this is much less important: observers routinely detected aircraft that were at times outside the range of the radar information, and much less frequently failed to note the presence of an aircraft in the monitored region of airspace. We look forward to completing data collection and generating a more thorough analysis of these data, including consideration of azimuth (for city lights, mountains, sun position, and the like) and elevation above the horizon.

References

Akerman A, Kinzly RE. (1979). Predicting aircraft detectability. *Human Factors, 21,* 277-291.

Brown JL. (1965). Flashblindness. *American Journal of Ophthalmology. 1965; 60,* 505-520.

Hood DC, & Finkelstein MA. (1986). Sensitivity to light. In Boff, KR, Kaufman, L, & Thomas, JP (Eds.), *Handbook of Perception and human performance.* New York: Wiley.

Editors. (1996). FAA, FDA put hold on outdoor lasers, but only at Las Vegas. *Air Line Pilot, 65,* 40-41.

O'Neal MR, Armstrong HG, Miller, RE. (1988). Further investigation of contrast sensitivity and visual acuity in pilot detection of aircraft. Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, OH: 1988; AAMRL-TR-88002.

Pulling, NH, Wolf, E, Vaillancourt, DR, & Dolliver, JJ (1980). Headlight glare resistance and driver age. *Human Factors, 22,* 103-112.

Olzak LA, Thomas JP. Seeing spatial patterns. In Boff, KR, Kaufman, L, & Thomas, JP (Eds.), *Handbook of Perception and human performance.* New York: Wiley.

Scott WB. (1995). Southwest pilot injured by laser. *Aviation Week & Space Technology,* Nov 20, 1995.

Sturgis, SP, & Osgood, DJ (1982). Effects of glare and background luminance on visual acuity and contrast sensitivity: Implications of driver night vision testing. *Human Factors, 24,* 347-360.

Thomas SR. (1994). Review of personnel susceptibility to lasers: Simulation in SIMNET-D for CTAS-2.0. Occupational and Environmental Health Directorate, Optical Radiation Division, Brooks AFB, TX: 1994; AL/OEO-TR-1994-0060.

Vos JJ. (1984). Disability glare--a state-of-the-art report. *C.I.E Journal, 3,* 39-53.

Weiss SA. (1996). Aviation officials look warily at scientific lasers. *Photonics Spectra,* May.

A STUDY OF DATA COMPRESSION BASED ON
HUMAN VISUAL PERCEPTION

Jer-Sen Chen

Assistant Professor

Department of Computer Science and Engineering

Wright State University

3640 Colonel Glenn Highway

Dayton, OH 45435

Final Report for:

Summer Faculty Research Program

Armstrong Laboratory

September 1996

A STUDY OF DATA COMPRESSION BASED ON
HUMAN VISUAL PERCEPTION

Jer-Sen Chen
Wright State University
Department of Computer Science and Eng.

Abstract

With the advances in hardware and software technology, multimedia that includes text, audio, imagery, and video is becoming part of our daily computing life. Even with faster computers, larger storage capacity, and higher communication bandwidth, data compression is still needed to overcome greater demand of multimedia data volume. This report presents a study on the various data compression techniques. including JPEG and MPEG, wavelet, and fractal image compression. In particular, issues related to compression schemes based on human visual perception are addressed. Research effort in the past in the are of human vision based compression techniques primarily focused on still images. A good example is the recommanded quantization matrix for discrete cosine transformation in JPEG and MPEG compression. Various quantization matrices have been proposed to accommodate different imagery as well as viewing conditions. Human vision models can also be incorporated in other image compression algorithms such as those using wavelets and fractals. Video compression based on human perception presents a great deal of challenge not only in terms of increased data volume but also in terms of more complicated spatial-temporal human vision model. This report also discusses possible study of vision based video compression techniques as well as available hardware technologies that facilitate the task.

A STUDY OF DATA COMPRESSION BASED ON
HUMAN VISUAL PERCEPTION

Jer-Sen Chen

Introduction

Advances in computing technology enable data content evolving from tradition simple text to multimedia presentation. The increasing in both processing speed and data transfer rate could not overcome the greater demand of sophisticated multimedia content. Among different types of media, digital video consumes more storage space and communication bandwidth than other types such as text and audio. For example, a typical video stream at 640x480 pixels resolution without any compression aid can result in more than 221 million bits per second if it runs at 30 frames per second and 24 bits per pixel true color mode. Even with current technology, especially the communication bandwidth, some form of compression is still very much in need.

A data stream consists of typically both entropy and redundancy. Entropy is the information content while redundancy is the part can be removed without loss of information. Lossless compression techniques are commonly used for text data files where exact bit level recovery is required. Lossless compression, and other compression techniques in general, are sensitive to noise and extra redundant information is needed to increase the integrity of the coding schemes during transmission.

Lossy compression techniques, usually applied to audio or video signals, can not recover exact bit level information. They are in general designed to exploit human perceptual capability such that losslessness can be achieved in terms of human perception. For example, coders and decoders are commonly designed to tolerate higher degree of distortion at frequency range where audio or video perception is much less sensitive. On the other hand, high fidelity is maintained at frequency range where perceptual sensitivity is high.

Still Image Compression

Image compression techniques usually explore spatial coherence of pixel data, that is, most pixels in the nearby area share common color values except at object boundaries. Redundancy occurs where individual pixel is coded among homogeneous region where pixel values are similar. To explore and remove redundancy in an image, simple spatial form of differentiation can be employed such as Adaptive Differential Pulse Code Modulation (ADPCM) or Run Length Encoding (RLE). A more

sophisticated scheme can explore the coherence in the frequency domain such as Discrete Cosine Transformation (DCT) used is JPEG, or subband coding and wavelets.

Predictive Coding

Predictive coding explores the correlation between the neighboring pixel values in spatial domain. The most common approach is Differential Pulse Code Modulation (DPCM) and its variant Adaptive Differential Pulse Code Modulation (ADPCM). Prediction of a pixel value is generally a linear combination of its neighboring pixel values. The prediction is subtracted from the original signal to form the differential signal which is then quantized and encoded. The deign of a DPCM scheme is therefore totally dependent on the optimization used in the predictor as well as the quantizer. Adaptive DPCM or ADPCM, instead of using fixed predictor and quantizer throughout an image, adjusts adaptively according to local image statistics. ADPCM is a very simple and straightforward coding scheme, both in compression and decompression. Its performance however is not so good as compared to other transformed domain coding scheme. When applying to regular photographic images such as the one shown in figure 1, artifacts start to emerge when the bit rate goes below 1 to 2 bits per pixel (bpp).

Figure 1: Original test image at 512x512x8 bits resolution

Transform Coding and JPEG

By far the most popular image compression scheme for photographs is the one proposed by Joint Photographer Expert Group (JPEG). Its algorithmic simplicity as well as its outstanding quality performance helps gain its popularity. The algorithm consists of, after decomposing the original image into eight by eight blocks, a Discrete Cosine transformation (DCT), a quantizer, and variable length encoding scheme such as Huffman codes. The decoding is exactly the reverse and makes JPEG a highly symmetric scheme, that is, compression and decompression take approximately same amount of computational cost.

The outstanding performance of JPEG compression is measured not only objectively but also subjectively. Discrete cosine transformation has long been recognized as almost optimal as KLT. It produces very high objective measurement such as signal to noise ratio (SNR). Also because its periodicity tapers off at boundaries, the artifact of blockiness is less pronounced compared to other transformation such as Discrete Fourier Transformation (DFT). Blockiness is always the most noticeable artifact in most block based compression schemes at low bit rate.

The design of quantization matrix contributes to another degree of subjective performance improvement in JPEG compression. The quantization matrix proposed by JPEG is the following:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Note that lowest frequency (DC or average greylevel) is located at upper left corner while the highest frequency is at lower right. After transformation, each DCT coefficient is divided (quantized) by the corresponding coefficient in the 8x8 quantization matrix. Therefore the larger the quantization coefficient, the more likely the DCT coefficient is quantized to zero. Figure 2 illustrates results of JPEG compression. In general, photographs digitized at 512 by 512 resolution can retain very high image quality when compressed above 0.5 bits per pixel.

Figure 2: JPEG compression at 0.5 bpp (top) and 0.2 bpp (bottom)

The quantization matrix is designed to account for human visual perception, such as low contrast sensitivity at high spatial frequencies. The numbers in the quantization matrix therefore, in general, increase from upper left to lower right though not strictly monotonic. Many researchers have been studying the relevance of the DCT quantization coefficients to human vision [Kline92, Watson94, Solomon94].

Subband and Wavelet Image Coding

Another class of image compression scheme decomposes original spatial data into different frequency band and orientation. The wavelet transformation decomposes visual information by scale or resolution, where maintain the locality at each scale. Quadrature mirror filters (QMFs) are usually used in subband encoding, followed by a subsampling process. Since each frequency band takes only a fraction of the whole spectrum (within Nyquist limit), down sampling can be employed before source encoding is applied. Because human visual system appears to employ multiscale processing in nature, applying multiresolution decomposition such as wavelet transform makes sense from perceptual point of view. Efficient algorithms for computing wavelet transformation help establishing subband coding as one of the major contender in image compression. [Mallat89, Adelson87]

Similar to transform coding, wavelet image coding consists of three steps, namely transformation, quantization, and entropy encoding. As in JPEG compression, the lossy part of the processing is the quantization where quality trades off bit rate. The decoding process is also similar to JPEG, just inverses the encoding processing steps. Both encoding and decoding process in wavelet compression are symmetric, that is they require approximately same amount of computational cost.

Unlike JPEG compression, wavelet image coding does not employ 8x8 block by block decomposition. Wavelet transformation starts with the whole image, and through down sampling to reduce the size of next stage processing. Since it process the entire image at each scale, it requires higher computational cost than JPEG. On the other hand, since it is not block based, its blockiness artifact is much less pronounced compared to JPEG. This is particularly true when low bit rate encoding is used, such as at 0.2 bits per pixel. Figure 3 illustrates results of wavelet image compression.

Figure 3: Wavelet compression at 0.5 bpp (top) and 0.2 bpp (bottom)

Fractal Image Compression

Fractal geometry is usually used in describing nature. Using fractal transforms to achieve image compression can be attributed largely by Michael Barnsley [Barnsley93]. The images to be encoded are usually quite different from mathematical fractals. The self-similarities are only explored in the *local* sense [Fisher95], that is, from one local part of the image to another local part of the same image to be encoded.

A key issue in the fractal image compression is to fully exploit the self-similarities in the images to be encoded. The self-similarity is usually translated into a matching or mapping process between the domain and the range, which are different partitions of the same input image.

To account for feasibility, the range is derived from partitioning the input image into non-overlapping blocks, while the domain consists overlapping blocks. The fractal image compression is then defined as, for each range block, try to find the best matching block from the domain pool. Variable size partition of the range is usually used to achieve higher compression ratio, as long as all blocks are non-overlapping. Typical block sizes used in a 512 by 512 image range from 16 by 16 down to 4 by 4. This approach of simultaneously decomposing a square block into 4 smaller square blocks is call quadtree decomposition. Other schemes that decompose in one dimension at a time (HV-decomposition) can be used to improve performance.

The mapping from a domain block to a range block is accompanied by the following transformations:

- a decimation process, say from 16 by 16 to 8 by 8,

- a scalar offset to the entire block to account for average brightness,

- a scaling factor to account for the contrast, and

- a rotation index to account for geometric orientation.

The shrinking of the domain block is necessary for convergence in the iterative decoding process. Decimation of every 2 by 2 block into single pixel is usually used for efficient computing. Other decimation factors can be used in order to obtain a better match at the expense of higher computational cost at the encoding stage and extra storage for decimation ratio information.

Mean square error between the range block pixels and the transformed domain block pixels is usually used to quantify the distortion or loss caused by the compression scheme. The encoding process of the fractal image compression is therefore described as matching between range and domain blocks

where the goal is to minimize the matching errors. Assume the decimated domain block contains n pixels with greyscales $d_1, d_2, \ldots, d_n$ and the range block contains $r_1, r_2, \ldots, r_n$. The square error E between the range block and the transformed domain block, under scaling factor $\mathbf{S}$ and offseting factor $\mathbf{O}$, is therefore:

$$E = \sum_{i=0}^{n-1} (\mathbf{S}d_i + \mathbf{O} - r_i)^2$$

In order to minimize the square error E above, the following partial derivatives are set to zero:

$$\frac{\partial E}{\partial \mathbf{S}} = \sum_{i=0}^{n-1} 2d_i(\mathbf{S}d_i + \mathbf{O} - r_i)$$

and

$$\frac{\partial E}{\partial \mathbf{O}} = \sum_{i=0}^{n-1} 2(\mathbf{S}d_i + \mathbf{O} - r_i)$$

Let

$$\Delta = \sum_{i=0}^{n-1} d_i^2 \times n - \left(\sum_{i=0}^{n-1} d_i\right)^2$$

and

$$\Delta_s = \sum_{i=0}^{n-1} d_i r_i \times n - \left(\sum_{i=0}^{n-1} r_i\right)^2$$

then we get the best scale:

$$\mathbf{S} = \frac{\Delta_s}{\Delta}$$

and the best offset:

$$\mathbf{O} = \frac{\sum_{i=0}^{n-1} r_i - \mathbf{S} \sum_{i=0}^{n-1} d_i}{n}$$

The fractal image compression can therefore described as: for each range block, find the best transformed matching block from the domain pool. The encoding process is therefore very computationally expensive. For practical implementation, instead of finding the optimal match for each range block, the searching process stops once a matching domain block, within predefined error threshold, is found. On the other hand if no domain blocks found within the error threshold subdivision if the range block is performed. Non-uniform partition of the range pool such as using a quadtree is usually used to improve the image quality [Fisher95]. Figure 4 illustrates results of fractal image compression.
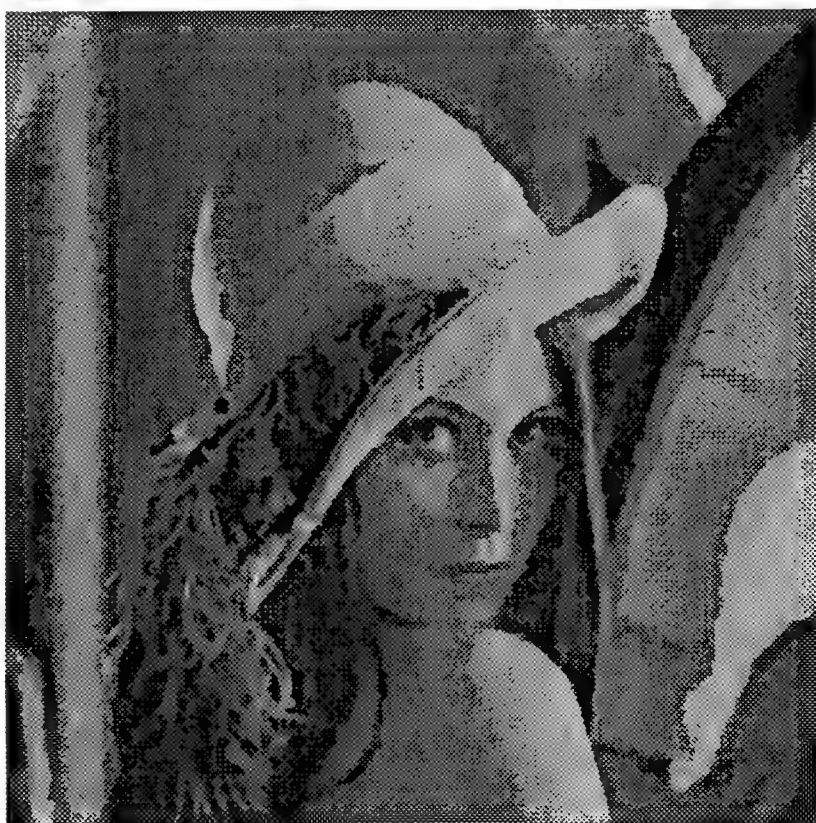
Figure 4: Fractal compression at 0.5 bpp (top) and 0.2 bpp (bottom)

Comparison of Still Image Compression Techniques

Root mean square error (RMSE) measure is commonly used to quantify the integrity of a coding scheme. It is formulated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(O_i - R_i)^2}{N}}$$

where O_i is the pixel value of the original image, R_i is the reconstructed pixel value, and N is the total number of pixels.

Comparing the above mentioned still image compression schemes, namely ADPCM, JPEG, wavelets, and fractals, all algorithms performs reasonably well when bit rate is above 1 bit per pixel (bpp). At 0.5 bits per pixel, all but ADPCM still maintain very good visual quality which is also reflected in their root mean square errors. While at lower bit rate, such as 0.2 bits per pixel, wavelet subband encoding slightly outperforms both JPEG and fractal compression, both in image visual quality and RMSE measure. The performance of fractal compression, however, can be improved by HV-decomposition than quadtree subdivision [Fisher95]. Also note that the computational complexity of the encoding part of fractal compression is much higher than the other two schemes.

At very low bit rate of 0.2 bits per pixel, blockiness emerges from both JPEG and fractal compression, though it is more pronounced in JPEG than in fractal compression. This undesirable visual artifact arises neither from discrete cosine transformation of JPEG nor from fractal transformation, rather, it is caused by the block based nature of their algorithmic implementation. Wavelet compression does not suffer the artifact because it processes the entire image as a whole at every scale and orientation.

Image Compression Based on Human Perception

The above root mean square error measure in general correlates well with human perceived difference between original image and compressed image. But counter examples can be easily devised that, in certain cases, same amount of RMSE actually translates into quite different perceived distortion [Daly93]. Alternative error measures can be utilized in designing an image compression technique to account for factors that are more closely related to factors human visual perception such as contrast sensitivity function and masking effect.

Contrast sensitivity, reciprocal of contrast threshold, is a function of spatial frequency. Noise or distortion arises from lossy image compression below the contrast threshold can be treated as visually lossless. JPEG compression takes into consideration of human visual perception in terms of non-uniform weights in DCT quantization matrix. Different coefficients in the matrix are assigned to

reflect different spatial frequency component weighting. The asymmetry of the matrix also reflects different spatial responses in horizontal and vertical orientation at same spatial frequency.

Weighted error measure in the frequency or transformed domain can also be designed to account for contrast sensitivity function in other compression scheme. Improvements in image quality was reported by applying weighted DCT domain error measure to fractal compression [Chen95] though it introduced additional computational complexity to already intensive algorithm. Here's the brief description of the approach.

An error measure instead of RMSE in the transformed domain, say discrete cosine transformation, is used to account for the dependency of human contrast sensitivity to spatial frequency. In particular, a weighted square error measure in the DCT domain is defined as:

$$E = \sum_{i=1}^{n-1} w_i^2 (\mathbf{S} D_i - R_i)^2$$

where D_i is the DCT coefficients of the decimated domain block d_i, R_i is the DCT coefficients of the range block r_i, and w_i are error weighting coefficients that account for the dependency of the human visual perception on spatial frequencies. The weights can be chosen as the quantization matrix recommanded by JPEG committee.

Note that the summation starts from 1 instead of 0, and there is no offset factor O in the error measure. This is because R_0 and D_0 are just the DC component of the range and domain blocks, we can therefore set brightness offset factor to equate the difference.

By setting the following partial derivative to zero:

$$\frac{\partial E}{\partial \mathbf{S}} = \sum_{i=1}^{n-1} 2 w_i^2 D_i (\mathbf{S} D_i - R_i)$$

we get the best scale:

$$\mathbf{S} = \frac{\Delta_s}{\Delta}$$

where,

$$\Delta = \sum_{i=1}^{n-1} w_i^2 D_i^2$$

and

$$\Delta_s = \sum_{i=1}^{n} w_i^2 D_i R_i$$

The best offset $\mathbf{O}$ is therefore simply the adjustment of the DC component:

$$\mathbf{O} = R_0 - \mathbf{S} \times D_0$$

Note that, for practical implementation, each DCT-transformed block is pre-multiplied by the weighting coefficients w_i.

Finally, improvements can also be made in visual performance by incorporating domain knowledge [Gray93]. For instance, intensity weighted compression can be used in medical imaging such as X-ray or MRI, as well as infrared thermal imaging. In such imagery, the high intensity areas signify important events and therefore are assigned with higher weights. Manually marked area of interest can also be used to assign weights to error measure. Notice that these techniques do not depend on viewing conditions, such as pixels per degree to reflect dependency of visual perception on spatial frequency.

From Image to Video Compression

The spatial acuity of human visual system depends on retinal image velocity as well as spatial frequency content of the scene. Temporal integration time of the visual system limits the temporal bandwidth. The reduction of spatial acuity in the presence of motion is usually interpreted using a rectangular "window of visibility" [Watson86] in two dimensional spatiotemporal frequency response. It has also been used to account for study of some artifacts caused by mismatch between the scene update rate and display refresh rate [Chen93].

The Nyquist sampling rate to avoid aliasing is at least twice the maximum of the signal frequency. In the temporal domain, the aliasing is usually observed in the presence of fast moving objects. A typical anti-aliasing technique involves certain degree of smoothing or averaging, either pre-sampling or post-sampling. And a typical approach for temporal anti-aliasing is using *motion blur* algorithm [Potmesil83]. When a camera films a fast moving object, motion blur is generated through the integration process during the time period when the shutter is open. The techniques of generating motion blur are often employed to computer generated animation to produce more natural and realistic video.

Still image compression techniques discussed in the previous section can mostly be extended to compression video signals with generalization of two dimensional processing to three dimensions adding the time dimension. This extension, of course, adds great deal of computational complexity to the encoding-decoding process. To reduce the complexity so that reasonable frame rate can be achieved, one can process video signals as collection of independent images and process each frame independently. Without exploring temporal coherence in video signals, very limited compression ratio can be achieved under normal quality requirement.

Wavelet or subband coding can be directly extend to video compression, either by using a three-dimensional subband decomposition or simply by processing video sequence frame by frame individually. Adaptive bit allocation involving the additional dimension, namely the temporal dimension, is a very difficult optimization problem. Perceptual optimization presents even a greater degree of difficulty. Bit allocation can based on energy and/or motion fidelity. Adaptive exchange of spatial and temporal resolution can also be used [Podilchuk90, Podilchuk91]. Similarly, fractal compression of video signals can either be done by matching three dimensional spatial-temporal blocks, or simply by matching two dimensional spatial blocks across the video sequence. Matching directly using three dimensional blocks increases computational cost tremendously and renders itself unfeasible even with current hardware technology.

MPEG Video Compression Standards

Even though human contrast sensitivity function behaves as a bandpass filter both spatially and temporally, compression techniques rarely are employed symmetrically along time dimension as in the other two spatial dimensions. Rather, temporal coherence is usually explored in terms of motion compensation. Motion compensation captures temporal coherence in video sequence and is usually implemented using certain square block sizes. Image blocks in the current frame are first predicted from adjacent blocks in the neighboring frames. This scheme is used in the Motion Picture Expert Group (MPEG) compression standard as well as other video encoding standards. Searching for a matching block can be very time consuming depending on algorithm and search range employed. This results in asymmetry between encoding and decoding complexity. That is, encoding takes higher computational cost than decoding since it has to detect and match block similarity in adjacent frames.

Matching functions can be designed to accommodate human visual perception. Commonly used matching functions used in motion estimation include: [Hang95]

- normalized cross-correlation function,

- mean square error,

- mean absolute difference, and

- number of thresholded difference.

Model Based Schemes

Data compression always aims at removing redundancy in a signal and therefore derive a compact descriptor of the signal content. The idea behind the model based coding approach is to represent the

scene content with a predefined model. For instance, for video telephone application, human facial models as well as animation dynamics of facial movement can be used for model based compression. Different applications may require different models.

Models can be two dimensional image based or three dimensional geometry based. For instance in the facial animation used in the video phone, two dimensional features such as contours and region can be used. Video sequences are usually encoded as contour or region change in both shape and position.

Three dimensional approaches, on the other hand, use full parameterized geometry models. The object geometry can be described using wire frame polygonal model, or parameterized using feature points of the model. Detection and recovery of model features using video camera is a very difficult problem and has been long researched in the image understanding research community. Some artificial constraints had been proposed to alleviated the difficulty in solving inverse problem, such as adding marking on the model or impose initial positional constraint on the model. These simplifications suffer the loss of naturalness in the practical operational environment.

Hardware for Video Compression

As video becomes an essential multimedia component, computers are now equipped with real time video processing capability. JPEG compression is a symmetric technique that the coder requires about same computational complexity as the decoder. MPEG compression inherits, in particular the intraframe encoding portion, substantial portion of JPEG compression components such as DCT and entropic coding. The interframe coding which employs motion compensation, however, creates asymmetry between coder and decoder. Real time requirement for the decoder of MPEG on a personal computer with hardware acceleration on a video board is becoming popular-these days. Encoding a raw video signal into a MPEG stream still requires special hardware to achieve real time performance. In other more complicated video compression techniques, even when real time can not be achieved, researchers can still take advantage of hardware acceleration to reduce time to develop research prototype or to derive experimental results.

When designing a video compression hardware, we need to take into account computing part as well as memory requirements, and the interconnection between all modules. The computing part can be improved by both algorithmic optimization and multiprocessing. Fine grain parallelism is commonly used in image or video processing. The memory requirements are influenced by the multiple access of original data and intermediate results. The interconnect network bandwidth depends on how frequent computing part communicates with memory modules.

Programmable hardware or multiprocessor systems are desirable for studying perception based video compression techniques. Digital signal processing (DSP) chips can be used to speed up well defined functionality such as DCT, wavelet transform, and even motion estimation. They can be used to facilitate human vision related study if the required processing is straightforward such as varying DCT quantization matrix to account for contrast sensitivity and masking effect. Field programmable gate array (FPGA) [Xilinx94] chips, however, can be used when greater flexibility is required for studying various types of vision based video compression performance.

Popular DSP chips such as TMS320C80 from Texas Instruments [Beinart94] consist of 4 DSP and 1 master RISC on-chip processors. With 50 MHz clock frequency that delivers over two billion RISC-like operations per second and I/O data rate of 400 million bytes per second, MPEG compression can be implemented using one or more TMS320C80 to achieve real time requirements. DSP chips provide ideal building blocks for a wide range of study in human perception based video compression techniques, including discrete cosine transformation, motion estimation, and wavelet and subband coding.

FPGA chips, on the other hand, provide flexibility at the expense of operative sophistication. Rapid advancements in FPGA technology have already push new FPGAs with more than 40,000 gates that can run at speeds of more than 50 MHz. Mapping video compression algorithms onto hardware architecture may be too complicated or expensive to be optimized or sometimes just to be correct. Some researchers have started applying FPGA chips in real time video compression techniques [Schoner95].

Conclusions

The importance of data compression can never be understated even with the advances in hardware and software technologies. Faster processors, higher networking bandwidth, and larger storage capacity prove only to invoke higher human demand of multimedia. Lossless compression techniques are used when applications require accuracy at single bit level. Lossy compression techniques, on the other hand, are usually used where the integrity of the information is required at a more relaxed level and losslessness can be achieved through application specific perceptual requirements. JPEG still image compression takes into account the human perception factors through the design of the quantization matrix. Wavelet or subband compression techniques, apprxoimate scale and orientation decomposition human perception process, and offer better image quality especially at very low bit rate compression. Fractal image compression presents an alternative to compress image that exploit

local self-similarities within the image.

It is essential to incorporate human visual perception in the design of a compression technique since the output eventually goes to human eyes. Many researchers have embarked on studying various perceptual issues in still image compression but very limited in video compression. Study of peception based video compression increases the complexity in both data volume and human perceptual factors. Video signals add additional temporal dimension in the data which in return increase the data volume usually by 2 to 3 orders of magnitude. The computational complexity increases at least by the same factor. Study of human perception of video signals not only introduces the additional temporal dimension, but also the interaction of visual perception in spatial-temporal domain. Some hardware acceleration techniques are definitely needed in conjunction with software algorithmic design to produce reasonably robust and timely results.

While vision researchers are still exploring human vision models, study of perception based video compression techniques must be flexible enough for both different algorithms as well as various vision models. A video sequence is treated as a three dimensional structure of pixels with two spatial dimensions and one temporal dimension. Research results in DCT quantization matrix design for JPEG compression can be readily extended to a three dimensional quantization matrix that accounts also for temporal frequency response. Extending from still image compression techniques, the transformation process can be three dimensional block based or can operates on entire video sequence such as in wavelet compression.

On the other hand, as in most video compression standards, temporal correlation can be exploited through motion compensation. Human visual perception can be incorporated in terms of the definition of block matching errors which renders the optimality in the matching process. How the design of spatial quantization matrix interacts with the choice of block matching error measure is another area to be explored.

Finally, direct generalization of fractal image compression into a three dimensional block similarity search process is definitely too computational expensive to be explored with present technology. Both hardware acceleration and software simplification are necessary components for studying human perception based fractal video compression.

Acknowledgement

References

[Adelson87] E. Adelson, E. Simoncelli, and R. Hingorani, "Orthogonal Pyramid Transforms for Image Coding," SPIE Visual Communication and Image Processing II, pp. 50-58, 1987.

[Barnsley93] M. Barnsley, Fractals Everywhere, 2nd edition, Academic Press, Boston, MA., 1993.

[Beinart94] , J.H. Beinart, "MVP: A Multimedia Video Processor Architecture," IEEE Workshop on VIsual Signal Processing and Communications, Rutgers, New Jersey, 1994.

[Chen93] , Jer-Sen Chen, "A Study on the Effects of Low Update Rate Visual Displays," Proceedings of the Society of Information Display '93, Seatle, Washington, May 1993.

[Chen95] , Jer-Sen Chen, "Fractal Image Compression Based on Visual Perception," Proceedings of Human Vision, Visual Perception, and Displays 1995, San Jose, California, February 1995.

[Daly93] , S. Daly, "The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity," in Digital Images and Human Vision edited by A.B. Watson, MIT Press, Cambridge, Masschusetts, 1993.

[Fisher95] Y. Fisher, Ed., "Fractal Image Compression, Theory and Application," Springer-Verlag, New York, 1995.

[Gray93] , R.M. Gray, P.C. Cosman, and K.L. Oehler, "Incorporating Visual Factors into Vector Quantizers for Image Compression," in Digital Images and Human Vision edited by A.B. Watson, MIT Press, Cambridge, Masschusetts, 1993.

[Hang95] H.M. Hang and Y.M. Chou, "Motion Estimation for Image Sequence Compression," in Handbook of Visual Communication, edited by H.M Hang and J.W. Woods, Academic Press, 1995.

[Kline92] S.A. Kline, A.D. Silverstein and T. Craney, "Relevance of human vision to JPEG-DCT compression," Human Vision, Visual Processing, and Digital Displays III, SPIE Vol. 1666, pp 200-215, SPIE, 1992.

[Mallat89] , S. Mallat, "Multifrequency Channel Decompositions of Images and Wavelet Models," IEEE Transactions on Acoustics, Speech, and Signal Processing, pp 2091-2110, 1989.

[Peitgen88] H. Peitgen and D. Saupe, Ed., The Science of Fractal Images, pp 100-101, Springer-Verlag, New York, 1988.

[Podilchuk90] C.I. Podilchuk and N. Farvadin, "Perceptually Based Low Bit Rate Image Coding," in Proceedings of ICASSP, 1990.

[Podilchuk91] C.I. Podilchuk, N.S. Jayant, and P. Noll, "Sparse Codebooks for Quantization of Non-dominant Sub-bands in Image Coding," in Proceedings of ICASSP, 1991.

[Potmesil83] M. Potmesil and I. Chakravarty, "Modeling Motion Blur in Computer-Generated Images," Computer Graphics, Vol. 17, No. 3, July 1983, pp. 389-399.

[Rabbani91] M. Rabbani and P.W. Jones, Digital Image Compression Techniques, SPIE Press, Bellingham, Washington, 1991.

[Schoner95] ,B. Schoner, J. Villasenor, S. Molloy, R. Jain "Techniques for FPGA Implementation of Video Compression." ACM/SIGBA International Symposium on Field-Programmable Gate Arrays, 1995.

[Solomon94] J.A. Solomon, A.B. Watson, and A.J. Ahumada, Jr., "Visibility of DCT Quantization Noise: Contrast Masking," SID 94 Digest, 1994.

[Watkinson95] J. Watkinson, "Compression in Video and Audio," Focal Press, an imprint of Butterworth-Heinemann Ltd., Linacre House, Jordan Hill, Oxford OX2 8DP, 1995.

[Watson86] A.B. Watson, A.J. Ahumada, Jr., and J.E. Farrell, "Window of Visibility: a Psychophysical Theory of Fidelity in Time-Sampled Visual Motion Displays," Journal of Optical Society America A, Vol. 3, No. 3, March 1986, pp. 300-307.

[Watson94] A.B. Watson, J.A. Solomon, and A.J. Ahumada, Jr., "Visibility of DCT Quantization Noise: Effects of Display Resolution," SID 94 Digest, 1994.

[Xilinx94] The Programmable Logic Data Book, Xilinx Inc., San Jose, California, 1994.

A SEQUENTIAL OPTIMIZATION ALGORITHM FOR PERSONNEL ASSIGNMENT BASED ON CUT-OFF PROFILES AND A REVISION OF THE BROGDEN TABLE

Cheng Cheng
Assistant Professor
Department of Mathematical Sciences

The Johns Hopkins University
Baltimore, MD 21218

A SEQUENTIAL OPTIMIZATION ALGORITHM FOR PERSONNEL ASSIGNMENT BASED ON CUT-OFF PROFILES AND A REVISION OF THE BROGDEN TABLE

Cheng Cheng
Assistant Professor
Department of Mathematical Sciences
Johns Hopkins University

Abstract

Two problems in the methodological research and practice of personnel psychology are considered. A sequential optimization algorithm flexible enough to incorporate various hard-to-quantify constraints on the personnel assignment solutions is developed and assessed by Monte Carlo simulation. The results show that the algorithm is able to produce solutions with allocation efficiency close to the theoretical ceiling. The open problem of establishing an accurate adjustment formula for the Brogden (1959) table used in estimating assignment benefits is solved by taking a distribution-theoretic approach.

A SEQUENTIAL OPTIMIZATION ALGORITHM FOR PERSONNEL ASSIGNMENT BASED ON CUT-OFF PROFILES AND A REVISION OF THE BROGDEN TABLE

Cheng Cheng

1. Introduction

The Armstrong Laboratory, the Army Research Institute for the Behavioral Sciences, the Navy Personnel Research and Development Center, and the Center of Naval Analysis co-sponsored a project that developed a "Join-Service Classification Research Roadmap" (Campbell & Russell, 1994). The report addressed various essential issues in personnel classification and job assignment procedures, including classification efficiency and its associated statistical measures such as the mean predicted performance (MPP) defined as the overall average predicted job performance scores of the applicants in an assignment solution (cf. Section 2). The report suggested that "Effective classification must consider many objectives and constraints to assign recruits to the job where they can perform effectively, while maintaining the efficiency of the recruiting and training system. Two major foci of recent classification research have been to develop classification methods that maximize MPP and to develop optimization methods to maximize classification objectives while satisfying constraints." This suggestion implies the need for a comprehensive sensitivity analysis of the effects of variation in critical features of a personnel system on the various assignment procedures to capture the potential gain, given constraints.

With the above points in mind, this study focuses on two important problems in the methodological research and practice of personnel psychology. For a long time there has been a need for an optimization algorithm alternative to the Linear Programming (LP) based procedures (Darby, Kyllonen, & Skinner, 1996). For various practical reasons, such as hard-to-quantify constraints on the assignment solutions, the algorithm should be flexible, in a sequential fashion, and coherently incorporate the information about the qualification cut-off scores of jobs (the cut-off score profile) often used in practice (Kyllonen & Darby, 1996). An associated open problem is how systematic changes in the cut-off score profile affect the allocation efficiency as measured by

the MPP. A sequential optimization algorithm for personnel assignment based on cut-off score profiles is developed in this study in an effort to accommodate the above considerations. Because of the simplicity in its core logic (cf. Figure 1 in Section 2), although not implemented yet, the algorithm is flexible enough to readily incorporate various constraints on the assignment solution, such as job priorities, applicants' job/regional preference, Affirmative Action considerations, job qualification criteria based on multiple scores, etc. This algorithm provides an alternative to the LP-based method using the transportation-problem model, in which the above-mentioned constraints have to be incorporated into a very complex payoff calculation model (Hendrix, Ward, Pina, & Haney, 1979). Monte Carlo simulation assessment shows that the algorithm is able to produce near-optimal solutions measured by the MPP, a result qualitatively consistent with the benchmark result in Grobman, Alley and Pettit (1995), although cut-off scores, rather than decision indices, are used here. In addition to sub-optimal personnel selection and assignment, the algorithm also generates/recommends a cut-off score profile that is geared toward the optimal assignment solution. Thus the algorithm also serves as a tool for systematically studying the inter-relationship among the cut-off score profile, the job quotas or rejection rate, and the MPP of assignment solutions.

There has been long and continued interest in the estimation of benefits of personnel assignment; see for example Brogden (1959), Rosse, Whetzel, & Peterson (1994), Alley & Darby (1994), and the references therein. Under three reasonable simplification assumptions enabling theoretical derivation, Brogden (1959) calculated a table (referred as the Brogden table hereafter) showing the theoretical MPP values in a multiple job system with or without rejection when the performance scores of the jobs are uncorrelated. For the more general case of equally correlated job performance scores Brogden (1959) proposed an adjustment formula. Using Monte Carlo simulation, Alley & Darby (1994) discovered that Brogden's adjustment greatly under estimates the actual MPP in the cases of non-zero rejection rates, and provided a remedy. It has been an open problem to establish sound and more accurate adjustment formula for the Brogden table. Taking a distribution-theoretic approach, this study has established an adjustment formula that gives more accurate results than the Alley-Darby formula.

Section 2 contains the description of the sequential optimization algorithm, together with illustrative examples. Section 3 contains the Monte Carlo simulation assessment of the algorithm. A revision to the Brogden table, together with the new adjustment formula, are given in Section 4. Section 5 contains some discussion and concluding remarks.

2. The Sequential Optimization Algorithm with Cut-off Profiles

This algorithm is designed to (1) find an assignment solution that maximizes the MPP represented by the overall average allocation efficiency (defined below), given the applicants data (the payoff matrix) and the jobs data including quota and a (initial) cut-off score profile; (2) produce a final (recommended) cut-off score profile that corresponds to the optimum assignment.

The algorithm takes as input (1) the job description data, including the number of jobs, quota on each job, and the (initial) cut-off score profile, and (2) the (predicted) performance scores of the applicants on each job (i.e. the payoff matrix). The algorithm produces as output (1) the solution to the assignment, together with per-job and overall total allocation and average allocation efficiency, (2) the final (recommended) cut-off score profile, and (3) possible remaining quota (i.e. unfilled jobs).

The total allocation efficiency of a job is the sum of the corresponding performance scores of the applicants assigned to that job; the average allocation efficiency of a job is its total allocation efficiency divided by the job quota; the overall total allocation efficiency is the sum of the total allocation efficiency of each job; the overall average allocation efficiency is the overall total allocation efficiency divided by the total quota. Formally, suppose there are m jobs with respective quotas $Q_1, Q_2, ..., Q_m$. Let P_{ij} be the performance score of the ith applicant in the jth job. Then the total allocation efficiency for the jth job is $T_j = \sum_{i \in A(j)} P_{ij}$, where $A(j)$ is the set of applicants assigned to the jth job; the average allocation efficiency for the jth job is $M_j = \left(T_j \big/ Q_j \right)$; the overall total allocation efficiency is $T = \sum_{j=1}^{m} T_j$; the overall average allocation efficiency (i.e. MPP) is $M = T \big/ Q$, with $Q = \sum_{j=1}^{m} Q_j$, the total quota.

Figure 1: The procedure of assigning one applicant

The algorithm processes the applicants one-by-one in a sequential fashion. For each applicant the algorithm first, in an effort to maximize individual performance, attempts to assign the applicant to the job corresponding to the highest performance score that is above the cut-off score (or scores) of the job. If the assignment is unsuccessful, the algorithm proceeds to make a series of decisions based on the current applicant's performance scores and the current assignment result. The detailed procedure is described in Figure 1.

The steps highlighted in Figure 1 are key steps for producing the (nearly) optimal assignment solution and a final suggested cut-off score profile corresponding to the optimal solution.

To better understand the algorithm, first consider two small-scale examples. An assessment of the algorithm using Monte Carlo simulation is given in the next section.

Example 1: Suppose that there are 4 jobs (job categories), labeled as M, A, G, E, with respective quotas 1, 2, 2, 1. The initial cut-off scores are set to -100000. This implies that every applicant is qualified for every job, which is in turn equivalent to using no initial cut-off scores. The jobs data file:
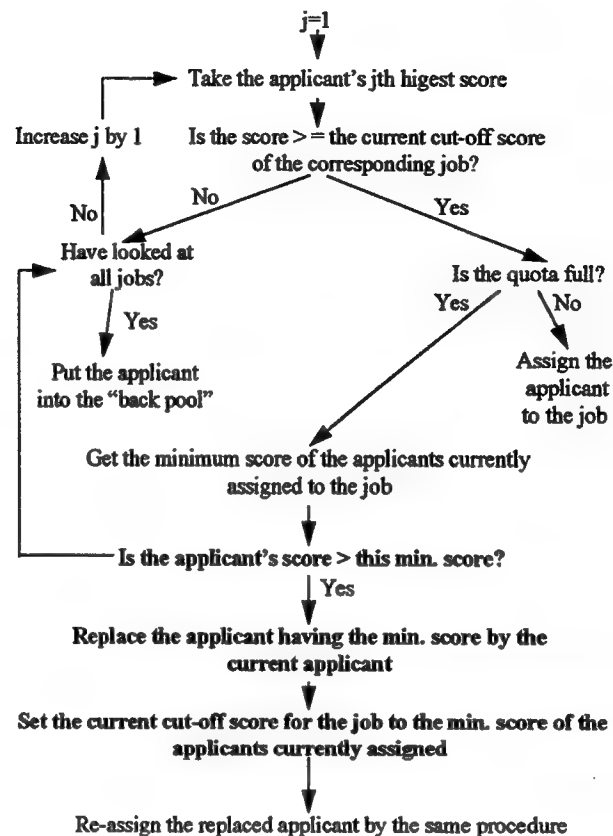
```
4                                        (number of jobs)
MAGE                                     (job symbols)
-100000. -100000. -100000. -100000.      (the initial cut-off score profile)
1 2 2 1                                  (quota of each job)
```

The applicants data file: Each row consists of the applicant's ID and performance scores; all rows together form the payoff matrix.

```
AA1  8  7  6  5
BB2  9  8  7  6
CC3  8  8  9  9
DD4  1  2 10  2
EE5  1  5  3  4
FF6 10  9  8  7
```

There are 6 applicants, so the assignment procedure in Figure 1 is performed 6 times:

Round 1: Applicant AA1 is assigned to Job 1.

Round 2: Applicant BB2 replaces AA1 in Job 1, and AA1 is re-assigned to Job 2.

Round 3: Applicant CC3 is assigned to Job 3.

Round 4: Applicant DD4 is assigned to Job 3.

Round 5: Applicant EE5 is assigned to Job 4.

Round 6: Applicant FF6 replaces BB2 in Job 1, and BB2 is re-assigned to Job 2.

The following assignment solution is then produced.

```
Job 1  M:
   FF 6  10.00                           (Applicant FF6 with score 10 is assigned to Job M.)
 ** Total Payoff for Job:  10.00
 ** Mean Payoff for Job:   10.00
 -- Final Cut-off Score:  10.00
 -- Remaining Quota: 0
Job 2  A:
   AA 1   7.00                           (Applicants AA1 and BB2 are assigned to Job A.)
   BB 2   8.00
 ** Total Payoff for Job:  15.00
 ** Mean Payoff for Job:   7.50
 -- Final Cut-off Score:   7.00
 -- Remaining Quota: 0
Job 3  G:
   CC 3   9.00                           (Applicants CC3 and DD4 are assigned to Job G.)
   DD 4  10.00
 ** Total Payoff for Job:  19.00
 ** Mean Payoff for Job:   9.50
 -- Final Cut-off Score:   9.00
 -- Remaining Quota: 0
```

Job 4 E:
   EE 5   4.00   (Applicant EE5 is assigned to Job E.)
 ** Total Payoff for Job: 4.00
 ** Mean Payoff for Job: 4.00
 -- Final Cut-off Score: 4.00
 -- Remaining Quota: 0

** Total Payoff: 48.00
** Mean Payoff: 8.00

     Example 2: There are 4 jobs (job categories), labeled as M, A, G, E, with respective

quota 1, 2, 2, 1. The initial cut-off score profile is 11.5, 9.5, 9.5, 8.

The jobs data:
4
MAGE
11.5 9.5 9.5 8
1 2 2 1

The applicants data:
AA1   8 7 6 5
BB2   9 8 7 6
CC3   8 8 9 9
DD4   1 2 10 2
EE5   1 5 3 4
FF6 10 9 8 7
GG7 5 4 3 2
HH8 1 10 1 9
II9   8 9 2 3
JJ10 12 6 10 7

The assignment solution:
Job 1 M:
   JJ 10 12.00
 ** Total Payoff for Job: 12.00
 ** Mean Payoff for Job: 12.00
 -- Final Cut-off Score: 12.00
 -- Remaining Quota: 0
Job 2 A:
   HH 8 10.00
 ** Total Payoff for Job: 10.00
 ** Mean Payoff for Job: 5.00
 -- Final Cut-off Score: 9.50
 -- Remaining Quota: 1   (The quota for this job is 2, but only 1 applicant is assigned.)
Job 3 G:
   DD 4 10.00
 ** Total Payoff for Job: 10.00
 ** Mean Payoff for Job: 5.00
 -- Final Cut-off Score: 9.50

```
   -- Remaining Quota: 1        (The quota for this job is 2, but only  1 applicant is assigned.)
Job 4  E:
  CC 3   9.00
 ** Total Payoff for Job:   9.00
 ** Mean Payoff for Job:   9.00
 -- Final Cut-off Score:   9.00
 -- Remaining Quota: 0


** Total Payoff:  41.00
** Mean Payoff:   6.83
```

Note that in this example, due to the quota limitation and the more stringent qualification standard represented by the initial cut-off score profile, the applicants AA1, BB2, EE5, FF6, GG7, and II9 are not assigned to any job, and the quotas of Jobs A and G are not completely filled.

Currently the algorithm puts the unassigned applicants into a "back pool", and no further action is taken. Moreover, the algorithm updates the cut-off score only when the job quota is full, otherwise the cut-off score retains its initial value given in the job description data file. In the situation when some job quota is not completely filled, it might be natural to lower the initial cut-off score by some amount; this action currently is not included in the algorithm. One possibility is to have a dialogue with the user at this point for a revision of the initial cut-off score profile, together with possible job priority information (e.g. some quota must be completely filled); then process the applicants in the back pool.

3. Monte Carlo Simulation Assessment of the Algorithm

3.1 The method. The overall average payoff M defined in the last section, representing the MPP of the assignment solution, is a natural indicator of the performance of the algorithm. One way to get a good sense about the algorithm's performance is to compare this quantity with its theoretical value under certain conditions. These conditions are generated by Monte Carlo simulation. The algorithm is assessed under the following assumptions: The applicants' predicted job performance scores follow the standard normal distribution, and the scores are uncorrelated across jobs. Under these conditions, the theoretical "floor" of the MPP is 0, which corresponds to completely random assignment without any rejection; the theoretical "ceilings" of the MPP are the

Brogden table entries (See Table 3 in Section 4 of this paper, Brogden 1959, and Alley & Darby 1994), which correspond to maximizing individual performances with or without rejection.

The following situational parameters are considered: the number of jobs, the rejection rate, the quota of each job, the initial cut-off score profile, and the number of applicants. To simulate the conditions under which the theoretical "ceilings" of the MPP are available, the jobs have to have the same quota (Brogden, 1959). The equal job quotas are determined as follows: Suppose there are m jobs and N applicants, and suppose the rejection rate is $100p$ percent. Then the approximate quota for each job is $(1-p)N/m$ rounded to the nearest integer. For example, if there are 4 jobs, 200 applicants, and the rejection rate is 0.4, then each job receives a quota 30. Quota rather than rejection rate is probably used more often in practice. In this simulation study the rejection rate is merely used for generating equal quota and comparisons with Brogden table entries.

Both the algorithm and the simulation are implemented in Turbo Pascal. The normal random variates are generated by the transformation method (Press, Teukolsky, Vetterling, & Flannery 1992, p. 289), and the uniform random numbers are generated by the Random function of Turbo Pascal.

3.2 The results. One thousand (1,000) simulation trials are run for several configurations of the situational parameters. In each simulation trial job scores of groups of 120 and 1000 applicants for 4 jobs, and groups of 280 and 2800 applicants for 7 jobs are generated and the assignment algorithm is run, producing one assignment solution and one (observed) value of the overall average allocation efficiency M for each configuration. So 1,000 values of the overall allocation efficiency are generated for each configuration. The mean of these 1,000 values is an estimate of the actual MPP, and is compared with the theoretical MPP ceiling from the revised Brogden table in Section 4. To get a better sense for the variation of the overall allocation efficiency M under the assumed conditions, the standard deviation of M, the 95% confidence interval of the actual MPP, and the five-number summary (the minimum, the first, second, and third quartiles, and the maximum) of M, are calculated based on the 1,000 simulated values.

For all the cases the initial cut-off scores are set to -100000, implying that every applicant is qualified for every job initially, which is in turn equivalent to using no initial cut-off score profile.

The results are summarized in Table 1 below. It is seen that the algorithm produces MPP values slightly lower than the theoretical MPP ceiling in all the experiments. The upper end of the 95% confidence interval is less than the theoretical ceiling in most of the cases, indicating the statistical significance that the algorithm's MPP does not achieve the theoretical ceiling. This statistical significance is primarily due to the small number of applicants. Recall that to obtain the theoretical MPP (ceiling), Brogden (1959) assumes infinitely many applicants. It can be clearly seen in Table 1 that as the number of applicants increases, the MPP estimate becomes closer to the theoretical MPP ceiling.

Table 1: *Simulation Summary. The underlined numbers are the theoretical MPP ceilings from the Revised Brogden Table in Section 4.*

configuration			rejection rate		
number of applicants	number of jobs	Statistics	0.0	0.5	0.9
120	4	theoretical	1.029	1.586	2.330
		mean	0.986	1.560	2.260
		std. dev.	0.066	0.080	0.147
		95%CI	(0.981,0.990)	(1.555,1.564)	(2.251,2.269)
		minimum	0.75	1.30	1.81
		1st quartile	0.94	1.51	2.16
		median	0.98	1.56	2.26
		3rd quartile	1.03	1.61	2.36
		maximum	1.20	1.81	2.81
1000	4	theoretical	1.029	1.586	2.330
		mean	1.017	1.583	2.321
		std. dev.	0.023	0.028	0.049
		95%CI	(1.016,1.019)	(1.581,1.585)	(2.318,2.324)
		minimum	0.93	1.49	2.18
		1st quartile	1.00	1.56	2.29
		median	1.02	1.58	2.32
		3rd quartile	1.03	1.60	2.35
		maximum	1.09	1.66	2.50
10000	4	theoretical	1.029	1.586	2.330
(100 simulation trials)		mean	1.026	1.585	2.329
		std. dev.	0.008	0.009	0.016
		95%CI	(1.025,1.028)	(1.583,1.587)	(2.326,2.332)
		minimum	1.00	1.57	2.29
		1st quartile	1.02	1.58	2.32
		median	1.03	1.58	2.33

Table 1. *(continued)*

number of applicants	number of jobs	Statistics	0.0	0.5	0.9
configuration			**rejection rate**		
		3rd quartile	1.03	1.59	2.34
		maximum	1.04	1.61	2.37
280	7	theoretical	<u>1.352</u>	<u>1.848</u>	<u>2.534</u>
		mean	1.316	1.830	2.484
		std. dev.	0.039	0.048	0.087
		95%CI	(1.314,1.318)	(1.827,1.833)	(2.475,2.489)
		minimum	1.20	1.67	2.21
		1st quartile	1.29	1.80	2.42
		median	1.32	1.83	2.4
		3rd quartile	1.34	1.86	2.54
		maximum	1.42	1.99	2.77
2800	7	theoretical	<u>1.352</u>	<u>1.848</u>	<u>2.534</u>
		mean	1.342	--	2.527
		std. dev.	0.012	--	0.028
		95%CI	(1.3416,1.343)	--	(2.526,2.529)
		minimum	1.29	--	2.44
		1st quartile	1.33	--	2.51
		median	1.34	--	2.53
		3rd quartile	1.35	--	2.55
		maximum	1.38	--	2.65

As mentioned earlier, the algorithm also serves as a tool for systematically studying the effect of the change in cut-off scores on the assignment performance. To demonstrate this point, a simulation is performed with several initial cut-off profiles. The results are summarized in Table 2. The relationship among the quota (or rejection rate), the (initial) cut-off score profile, and the MPP is clearly demonstrated: When the selection criterion (the initial cut-off score profile) is low, one obtains the near-optimal MPP. When the selection becomes more stringent but the quota remains large, for example, the case where the rejection rate is 0.5 and the initial cut-off score profile is [1.48, 1.48, 1.48, 1.48], the quota cannot be completely filled, producing an MPP=0.971, way below the theoretical ceiling; this represents the penalty for not filling all the quota -- here the MPP is represented by the overall allocation efficiency M, and recall that M is the total performance scores of the <u>assigned</u> applicants divided by the <u>total quota</u>. However, when the quota is small (or the rejection rate is high) relative to the size of the applicant pool, the MPP is less influenced by the tightening of the selection criterion -- note that the near-optimal MPP for

rejection rate 0.9 will not be affected until the initial cut-off scores become greater than 1.95. This might be a reflection of Taylor & Russell's (1939) argument from a different angle.

Table 2: *Simulation summary for cases with initial cut-off score profile. Number of applicants = 360; number of jobs = 4; number of simulations = 1000. Each simulation trial produces a final (recommended) cut-off profile, a remaining quota of each job, and an overall mean average performance value (a value of M). The "avg. recommended cut-off prof." and the "avg. remaining quota" are respectively the average recommended cut-off scores and the average remaining quota across the 1,000 simulations. the "avg. mean payoff" is the average of the 1,000 simulated M values, giving an estimate of the actual MPP. The underlined number in parenthesis is the theoretical MPP ceiling. The "std. dev.", "95% CI" and "5-number summary" are defined the same as those in Table 1.*

Statistics	Rejection Rate	
	0.5	0.9
quota	45, 45, 45, 45	9, 9, 9, 9
init. cutoff prof.	0.000, 0.000, 0.000, 0.000	0.000, 0.000, 0.000, 0.000
avg. recommended cutoff	1.008, 1.001, 1.008, 1.004	1.963, 1.958, 1.961, 1.962
avg. remaining quota	0.00, 0.00, 0.00, 0.00	0.00, 0.00, 0.00, 0.00
avg. mean payoff	1.576 (1.586)	2.303 (2.330)
std. dev. (mean payoff)	0.046	0.084
95% CI (mean payoff)	(1.573, 1.578)	(2.298, 2.308)
5-number summary:	1.44, 1.55, 1.57, 1.61, 1.73	2.03, 2.25, 2.30, 2.36, 2.62
quota	45, 45, 45, 45	9, 9, 9, 9
init. cutoff prof.	1.000, 1.000, 1.000, 1.000	1.000, 1.000, 1.000, 1.000
avg. recommended cutoff	1.051, 1.047, 1.051, 1.051	1.963, 1.958, 1.961, 1.962
avg. remaining quota	2.44, 2.34, 2.42, 2.43	0.00, 0.00, 0.00, 0.00
avg. mean payoff	1.528 (1.586)	2.303 (2.330)
std. dev. (mean payoff)	0.068	0.084
95% CI (mean payoff)	(1.524, 1.532)	(2.298, 2.308)
5-number summary:	1.29, 1.48, 1.53, 1.58, 1.74	2.03, 2.25, 2.30, 2.36, 2.62
quota	45, 45, 45, 45	9, 9, 9, 9
init. cutoff prof.	1.480, 1.480, 1.480, 1.480	1.480, 1.480, 1.480, 1.480
avg. recommended cutoff	1.480, 1.480, 1.480, 1.480	1.963, 1.958, 1.961, 1.962
avg. remaining quota	22.54, 22.65, 22.41, 22.45	0.00, 0.00, 0.00, 0.00
avg. mean payoff	0.971 (1.586)	2.303 (2.330)
std. dev. (mean payoff)	0.089	0.084
95% CI (mean payoff)	(0.965, 0.976)	(2.298, 2.308)
5-number summary:	0.71, 0.91, 0.97, 1.03, 1.25	2.03, 2.25, 2.30, 2.36, 2.62

4. A Revision of the Brogden Table

The mean payoff calculation problem studied by Brogden (1959) admits a natural formulation in terms of certain probability distributions. Technical details are left out to conserve space, and they are available from the current author or in a technical report at AL/HRM.

Brogden's assumptions enabling the theoretical calculation are equivalent to the following probabilistic model. Suppose there are m jobs. The applicant population represented by the

predicted performance scores can be treated as an m-component random vector $[C_1, C_2, \cdots, C_m]$, where C_j is the performance score of a (randomly selected) applicant on the jth job.

In the basic situation of unit validity ($R=1$) and zero correlation between the performance scores ($r=0$), each C_j is assumed to follow the standard normal distribution, and all C_j's are independent. Because each applicant is assigned to the job corresponding to the highest score, the allocation efficiency is given by the maximum component in the vector, i.e.

$$C^* = \max(C_1, C_2, \cdots, C_m).$$

In statistical terminology, C^* is the largest order statistic out of the m random variables. The mean allocation efficiency, or the mean predicted performance (MPP) after assignment, when the rejection rate $p=0$, is simply the expected value (or mean) of the largest order statistic C^*. By a known fact of the largest order statistic (See e.g. Larson, 1982, pp. 318-319) and the mathematical definition of mean, the theoretical MPP is given by the following integral.

$$(1) \qquad MPP = \int_{-\infty}^{\infty} xm[\Phi(x)]^{m-1} \phi(x)dx,$$

where $\phi(x) = (2\pi)^{-1/2} \exp\left(x^{-2}/2\right)$ is the standard normal probability density function (pdf), and

$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt$ is the standard normal cumulative distribution function (cdf). The entries on the second column of Table 1 in Brogden (1959) and Table A1 in Alley & Darby (1994) are the approximated values of the integral (1) for different numbers of jobs. When the rejection rate $p \geq 0$, the theoretical MPP is the following integral:

$$(2) \qquad MPP = \frac{1}{1-p} \int_{L}^{\infty} xm[\Phi(x)]^{m-1} \phi(x)dx,$$

where L is the $\left(p^{1/m}\right)$-th percentile (quantile) of the standard normal distribution.

In the more general situation where the validity $0 < R \leq 1$, and the performance scores $C_1, C_2, \cdots, C_m$ are equally correlated with pairwise correlation $0 < r < 1$, Brogden (1959) proposed a model in which all the m scores are related to a common variable K, plus an independent random perturbation following the normal distribution. Symbolically,

$$C_j = K + C'_j, \qquad j = 1,2,...,m,$$

where K follows the normal distribution with mean zero and variance $R^2 r$, and $C'_1, C'_2, \cdots, C'_m$ are independent and identically distributed by the normal distribution with mean zero and variance $R^2(1-r)$, and are independent of K. This model does provide the desired theoretical distribution of the performance scores, but in this situation a simple integral representation of the theoretical MPP like equation (1) or (2) no longer exists. For practice, a manageable adjustment for the table entries is desirable. The adjustment formula given by Brogden (1959) is correct only when the rejection rate is zero. Much improvement is obtained by Alley & Darby (1994) based on empirical studies using Monte Carlo simulation. With the known distributional properties of the performance scores and the properties of the quantile function (the inverse function of the cdf), an approximation formula can be established for adjusting the table values to a much better accuracy. This formula is given below together with a revised Brogden table.

With present-day computing power, the integrals (1) and (2), together with the quantity L in (2), can be easily calculated with numerical routines. The integrals can be calculated to a satisfactory accuracy with, say, a trapezoidal rule. Functions for calculating L are readily available in popular statistical packages such as SAS and S-plus.

Table 3 is a revised Brogden table computed using S-plus. Note that a new column V is added. Here m denotes the number of jobs. The table entries are the mean performance values for different number of jobs and rejection rates when the validity $R = 1$ and the between-job correlation $r = 0$. The quantity V is used for adjustment when $R \neq 1$ and/or $r \neq 0$. The adjustment formula is

$$R\sqrt{1-r}M_0 + \frac{R}{V}\sqrt{V^2(1-r)+r}\left(M_p - M_0\right),$$

where M_0 is the table entry for the given number of jobs and the rejection rate $p = 0$, M_p is the table entry for the given number of jobs and the rejection rate p. For example, for $R = 1$, $r = 0.85$, 16 jobs, and rejection rate $p = 0.6$, the adjusted mean performance is

$$\sqrt{1-0.85}(1.766) + \frac{1}{0.5432}\sqrt{(0.5432)^2(1-0.85) + 0.85}(2.295 - 1.766) = 1.604 .$$

Table 3: *A Revised Brogden Table*

m		Rejection Rate (percentage)									
	V	0	10	20	30	40	50	60	70	80	90
1	1.0000	0.000	0.195	0.350	0.497	0.644	0.798	0.966	1.159	1.400	1.755
2	0.8257	0.564	0.721	0.847	0.969	1.092	1.221	1.365	1.531	1.741	2.057
3	0.7480	0.846	0.986	1.100	1.210	1.322	1.441	1.573	1.727	1.923	2.220
4	0.7012	1.029	1.159	1.265	1.368	1.473	1.586	1.711	1.858	2.045	2.330
5	0.6690	1.163	1.285	1.386	1.485	1.586	1.693	1.813	1.955	2.136	2.413
6	0.6449	1.267	1.384	1.481	1.576	1.674	1.778	1.894	2.032	2.208	2.479
7	0.6260	1.352	1.465	1.559	1.651	1.746	1.848	1.961	2.096	2.268	2.534
8	0.6107	1.424	1.533	1.625	1.715	1.807	1.907	2.018	2.149	2.319	2.580
9	0.5978	1.485	1.592	1.682	1.769	1.860	1.957	2.067	2.196	2.363	2.621
10	0.5868	1.539	1.643	1.731	1.817	1.906	2.002	2.110	2.237	2.402	2.657
11	0.5773	1.586	1.689	1.775	1.860	1.948	2.042	2.148	2.274	2.437	2.689
12	0.5690	1.629	1.730	1.815	1.899	1.985	2.078	2.183	2.307	2.469	2.718
13	0.5614	1.668	1.767	1.851	1.933	2.019	2.111	2.214	2.338	2.497	2.745
14	0.5547	1.703	1.801	1.884	1.965	2.050	2.141	2.243	2.365	2.524	2.769
15	0.5487	1.736	1.832	1.914	1.995	2.078	2.168	2.270	2.391	2.548	2.792
16	0.5432	1.766	1.861	1.942	2.022	2.105	2.194	2.295	2.415	2.570	2.813
17	0.5381	1.794	1.888	1.968	2.047	2.129	2.218	2.318	2.437	2.592	2.832
18	0.5334	1.820	1.913	1.993	2.071	2.152	2.240	2.339	2.458	2.611	2.850
19	0.5291	1.844	1.937	2.015	2.093	2.174	2.261	2.360	2.477	2.630	2.868
20	0.5251	1.867	1.959	2.037	2.114	2.194	2.281	2.379	2.495	2.647	2.884
100	0.4294	2.508	2.580	2.642	2.705	2.771	2.843	2.925	3.024	3.154	3.360
200	0.4009	2.746	2.813	2.871	2.929	2.991	3.058	3.136	3.229	3.352	3.547
300	0.3865	2.878	2.941	2.998	3.054	3.113	3.179	3.253	3.344	3.463	3.653
400	0.3772	2.968	3.030	3.085	3.140	3.198	3.261	3.334	3.423	3.540	3.727
500	0.3704	3.037	3.097	3.151	3.205	3.262	3.324	3.396	3.483	3.599	3.783

The revised adjustment procedure is compared with those of Brogden (1959) and Alley & Darby (1994). The results are given in Table 4.

Table 4: *Comparison of Adjustment Procedures for* $R = 1$. *In the Empirical column are estimates of the actual MPP obtained by simulation in Alley & Darby (1994). Adjustment formulas should produce values close to these estimates.*

# Jobs	reject. rate	r	Empirical	Brogden	Alley-Darby	Revised
2	0.4	.45	1.00	.81	.97	.999
2	0.4	.85	.85	.42	.81	.842
2	0.6	.45	1.28	1.02	1.27	1.299
2	0.6	.85	1.16	.53	1.12	1.165
4	0.4	.45	1.30	1.10	1.25	1.301
4	0.4	.85	1.02	.57	.97	1.001

Table 4. (continued)

# Jobs	reject. rate	r	Empirical	Brogden	Alley-Darby	Revised
4	0.6	.45	1.58	1.27	1.52	1.588
4	0.6	.85	1.32	.66	1.26	1.333
8	0.4	.45	1.57	1.34	1.51	1.564
8	0.4	.85	1.17	.70	1.09	1.149
8	0.6	.45	1.84	1.50	1.76	1.843
8	0.6	.85	1.48	.79	1.38	1.478
16	0.4	.45	1.82	1.56	1.73	1.798
16	0.4	.85	1.29	.82	1.21	1.274
16	0.6	.45	2.06	1.71	1.96	2.071
16	0.6	.85	1.59	.89	1.49	1.604

In addition to the empirical comparison, it is instructive to compare the revised adjustment expression with those of Brogden (1959) and Alley & Darby (1994). First note that when the rejection rate p is 0, the second term of the revised adjustment formula is zero because $M_p = M_0$; the first term, $R\sqrt{1-r}M_0$, is nothing but Brogden's formula for $p=0$; moreover, this expression coincide with the Alley-Darby adjustment which is given in different notation; so all three adjustment formulas agree in this case. When the rejection rate is not zero however, the appropriate adjustment is not simply $R\sqrt{1-r}M_p$ as proposed by Brogden. Alley & Darby (1994) suggest adding the Brogden adjustment of the difference between m job performance (selection and classification) and single job performance (pure selection) to the single job performance. The revised adjustment formula states that the adjustment is the Brogden adjustment in the zero rejection case (pure classification) plus the properly scaled difference between the m job performance (selection and classification) and the pure classification performance.

5. Discussion and Concluding Remarks

Regarding the sequential optimization algorithm, the sequential processing of the applicants provides much flexibility for incorporating various constraints on the assignment that are not easy to quantify for Linear Programming (LP) based optimization algorithms. Such constraints may include, and certainly not limited to, job priorities, applicants' job and/or regional preferences, Affirmative Action considerations, etc. Alternative to building complicated payoff

calculation models to incorporate these constraints (Hendrix, Ward, Pena & Haney, 1979), these constraints can be directly inserted into the assignment decision logic described in Figure 1.

The flexibility of this sequential assignment algorithm also allows for assignment decisions for one or more jobs based on multiple test or performance scores. For example, some jobs may require the applicant to score above the cut-off on 2 or 3 tests, some jobs may require the applicant to score above the cut-off on one or another test. Again this type of constraint cannot be readily incorporated into LP-type procedures, but can be easily inserted into the decision logic in Figure 1.

The Monte Carlo simulation results show that under proper conditions the allocation efficiency of the algorithm is close to the theoretical ceilings. Because of the similarity in the basic assignment strategy, the simulation results here are qualitatively consistent with the benchmark results in Grobman, Alley and Pettit (1995).

Strictly speaking, the current algorithm is batch-sequential than pure sequential. It would be interesting to see how much MPP reduction might suffer from a pure sequential assignment, given the cut-off score profile, quotas, and other possible constraints.

In reality where both job requests and applicants come sequentially, the algorithm may be used during time periods (daily or weekly say) when several (or a set of) applicants and a set of job requests or class seats are available. There are at least three means to establish the initial cut-off score profile: use no initial cut-off scores, use cut-off scores obtained from the assignment solution by some other optimization procedure such as LP, and use the cut-off scores established earlier or in the previous time period by any means -experience, data analysis, subjective decision, some other optimization, etc. The third option is particularly useful for exploring the relationship among the cut-off score profile, the job quotas, and the allocation efficiency. Comparisons of the cut-off score profile generated by the algorithm with that established earlier will help in revising the cut-off scores for the current and future assignment tasks. Quantities such as job quotas, cut-off scores used, the overall average allocation efficiency, the cut-off scores recommended by the algorithm, etc. in various time periods are informative time series data for systematically studying the effect of cut-off scores on allocation efficiency.

The current algorithm is just a tip of the iceberg; many issues remain open. There is still a way to go before the algorithm becomes a genuine application software for personnel assignment and allocation. Much work remains for future research efforts.

The distributional approach to the Brogden table reveals some insights into the theoretical MPP calculation, and consequently a much more accurate adjustment formula is established. It is possible to extend this approach to other theoretical measurement calculations in personnel and education psychology, such as extending the Taylor-Russell tables (Taylor & Russell 1939) to the situation of multiple equally correlated jobs. Work along this line constitutes further research efforts.

References

Alley, W. E. and Darby, M. M. (1994) Estimating the benefits of personnel selection and classification: an extension of the Brogden table. *Armstrong Laboratory, AL/HR-TP-1994-0010*, Human Resources Directorate, Brooks AFB, TX.

Brogden, H. E. (1959) Efficiency of classification as a function of number of jobs, per cent rejection, and the validity and intercorrelation of job performance estimates. *Educational and Psychological Measurement.* **19**, 181-190.

Campbell, J. P. and Russell, T. L. (1994) Building a joint-service classification research roadmap: methodological issues in selection and classification. *Armstrong Laboratory, AL/HR-TR-1994-0013*, Human Resources Directorate, Brooks AFB.

Darby, M. M., Kyllonen, P. C. and Skinner, M. J. (1996) Personal conversations. *Armstrong Laboratory*, Human Resources Directorate, Brooks AFB, TX.

Grobman, J. H., Alley, W. E., and Pettit, R. S. (1995) The optimality of sequential personnel assignments using a decision index. *Armstrong Laboratory, AL/HR-TR-1995-0025*, Human Resources Directorate, Brooks AFB, TX.

Kyllonen, P. C. and Darby, M. M. (1996) Personal conversations. *Armstrong Laboratory*, Human Resources Directorate, Brooks AFB, TX.

Larson, H. J. (1982) *Introduction to Probability Theory and Statistical Inference* (3rd ed.) John Wiley and Sons, New York.

Press, W. H., Teukolsky, S. A. Vetterling, W. T. and Flannery, B. P. (1992) *Numerical Recipes in C.* Cambridge University Press, New York.

Rosse, R. L., Whetzel, D. L. and Peterson, N. G. (1994) Assessment of the classification efficiency of   selection/assignment systems. *Final Technical report for the period Nov. 1991 to April 1994, Contract Number F33615-91-C-0016.*

Taylor, H. C. and Russell, J. T. (1939) The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables. *Journal of Applied Psychology*, **23**, 565-578.

Tippett, L. H. C. (1925) On the extreme individuals and the range of samples taken from a normal population. *Biometrika*, **17**, 364-387.

Acknowledgment

PERCEPTUAL ISSUES IN VIRTUAL ENVIRONMENTS AND OTHER SIMULATED DISPLAYS

Elizabeth Thorpe Davis
Associate Professor

School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332-0170

PERCEPTUAL ISSUES IN VIRTUAL ENVIRONMENTS AND OTHER SIMULATED DISPLAYS

Elizabeth Thorpe Davis
Associate Professor
School of Psychology
Georgia Institute of Technology

Abstract

Virtual environments are multisensory and highly interactive display systems that come in a myriad of flavors and varieties. These VE systems can serve a multitude of purposes within the scientific, medical, military, industrial, and entertainment fields in ways that more traditional human-computer interfaces simply cannot. Because of all the potential and actual uses for VE systems, developing an optimal VE system is a high priority.

Developing an optimal VE system requires knowing and capitalizing on the capabilities and limitations of human perception, both within a given sensory modality and integrated across sensory modalities. Yet, no available VE system can fully exploit the capabilities of human perception, especially those of human vision. These technological limitations can impose some perceptual tradeoffs in utilizing available VE systems that one must carefully consider.

Conversely, VE systems provide an opportunity to answer some fundamental questions about how humans build up percepts about what is out there and what is going on, both within a given sensory modality (e.g., vision) and integrated across sensory modalities. Although VE systems ideally could mimic real-world experiences, they are not bound by the limitations of the real world (e.g., gravity and the laws of physics). Thus, perception in the simulated world of a VE system can dramatically differ from that in the real world. That is, VE systems allow us to test some limitations and capabilities of human perception in ways that more traditional displays and the real environment do not.

Our challenges this summer were to tackle these issues by thinking and reading, by setting up and conducting some pilot studies to explore the formation of multistable percepts within a virtual environment, and by writing a draft of a review paper based on these ruminations and preliminary results.

PERCEPTUAL ISSUES IN VIRTUAL ENVIRONMENTS
AND OTHER SIMULATED DISPLAYS

There has been alot of hype in the media about virtual environments and cyberspace. Although there are innumerable definitions of a virtual environment, Kalawsky (1993) defines it as a synthetic sensory experience that communicates physical and abstract components to a human user or participant. The basic components of a VE system consist of a system for generating the virtual environment, a visual display, an auditory display, possibly a tactile display, a system for tracking head and hand as well as possibly tracking eye or body positions, a manipulandum (e.g., a dataglove or 3-D joystick) and/or a speech recognition interface for communicating with the virtual environment. But, the actual components of the VE system may vary, as may the specific implementations of each component.

A myriad of actual and of potential uses exist for virtual environments -- within scientific and medical areas, within both the military and industrial arenas, and for use in the field of entertainment. Virtual environments can be used for training, as in flight training and simulation, driver training and assessment, surgical training, and training astronauts to deal with nonterrestrial environmental conditions. They can aid medical doctors in actual surgical operations and scientists in visualizing spatial relationships and how different parts interact, as in the visualization of planets or the molecular dockings of atoms. Virtual environments can be used in the teleoperation of robots through hazardous or unknown terrains -- as in locating and removing a bomb in a building or in exploring Jupiter's surface. Architects can walk their clients through a virtual building, so that costly modifications and adjustments can be avoided. Virtual environments can be used for sheer entertainment, too. For instance, one could take a virtual vacation to the location of one's dreams sans air travel or explore the ruins of Mesa Verde as they appeared to inhabitants almost a thousand years ago. These virtual environment pursuits places various demands both on the hardware system and the human perceptual and cognitive system that processes these synthetic inputs.

How do we actually build up our percepts of what is out there and of what is going on in either a real or a virtual environment? Although most virtual environments emphasize visual input, experiences within a virtual environment usually are multisensory and highly interactive -- as they are in real environments. Yet, perception in the simulated world of a virtual environment can differ from real

world experiences. Some of these differences are intentional and some are coincidental. What are some of the perceptual tradeoffs and limitations when interacting with a virtual environment? What future challenges face the designers and developers of virtual environments?

These are questions we posed and began investigating this summer. The results of our ten week tenure at Wright Patterson Air Force Base include thinking and reading about these issues, setting up and conducting some pilot studies to explore the formation of percepts within virtual environments, as well as drafting a review paper for publication based on some of these ruminations and preliminary results.

How perception in the simulated world of VE differs from perception in the real world

Virtual environments often try to mimic the real-world environment. Yet, the very nature of virtual environments suggests we can simulate information not readily available to our senses in real-world environments and that we can interact with this virtual environment in ways that normally are prevented by real-world constraints. For example, a VE could allow us to defy gravity by floating through space (as in microgravity conditions) or walking on ceilings. A VE could also visually (or auditorally) present distance information from a laser range finder. It could present infrared or thermal images to aid the identification of objects viewed under low ambient luminance conditions. In short, a VE system can overcome some of the limitations of the human perceptual system and add to the store of information normally available.

A VE system also can eliminate perceptual information normally available to our senses in the real world (e.g., eliminating the 3D visual information normally provided by stereopsis or the kinesthetic feedback from touching and pushing against objects). Sometimes information may be eliminated on purpose because the additional information is not *cost-effective* and/or adds no useful information for the tasks performed. That is, sometimes information is eliminated or included by purposeful design to enhance perception in and interaction with virtual environments. But, sometimes information is eliminated because of current technological limitations in the available hardware and software.

In any case, there are always tradeoffs -- enhancing one type of information or capability may adversely affect the presentation and availability of other

information and capabilities. On one hand, accessing more information can overload the human or machine system (or both!) thus causing information processing to become resource limited and non-optimal (Norman & Bobrow, 1975; Stokes, Wickens, & Kite, 1990). On the other hand, severely degraded displays that greatly restrict available information can cause data-limited deficits in information processing that also are not optimal.

Some perceptual tradeoffs and limitations of perceiving and of interacting within a simulated world

Although virtual environment systems come in many flavors and varieties and have countless potential uses, all VE systems face challenges in their design, development, and implementation. For example, all VE systems currently lack the full fidelity of real-world experiences and many of them do not incorporate certain sensory inputs (e.g., haptic, tactile and force feedback as well as taste or smell).[1] Some of the challenges faced are similar to those encountered with more traditional HCI systems -- including flight simulators -- and, thus, VE design and development can benefit from some of the previous lessons learned.

One top priority is meshing the human-computer interface design of a VE system so that it exploits the capabilities and limitations of the human perceptual system as well as those of the virtual environment. Moreover, designers and developers are working on incorporating some of the missing features into VE systems, but hopefully only when they are deemed cost-effective and can enhance actual task performance. Although several pressing problems face designers and developers of virtual environment systems, not even the users of virtual environments agree which problem is the most pressing and critical. Several candidate perceptual problems in vision and audition are discussed here.

A. Some visual considerations

Some feel the most challenging problems within VE systems may be visual, given that humans depend so heavily on visual information about their real-world environments and that at least 50% of the human brain is involved in processing visual information.

---

[1]Sensorama is an early virtual environment system that actually did include olfactory information from various smells as the user traversed the virtual roadway (Kalawsky, 1993; Rheingold, 1991).

Field of view (FOV) vs. spatial resolution tradeoff. Of these visual problems in VE, one of the most pressing is to simultaneously provide the user with a wide field of view *and* with good spatial resolution. No existing virtual environment yet achieves the wide field of view of the human visual system operating in a real environment: The human's instantaneous binocular field of view is approximately 200 deg in the horizontal direction and 120 degrees in the vertical direction. But, to create an immersive visual environment, designers want a very wide field of view (wFOV). Unfortunately, such a wide field of view often leads to worse spatial resolution within the visual image, so that spatial details of objects are lost and curved or diagonal edges appear jagged (i.e., spatial aliasing occurs). This is the well-known FOV versus spatial resolution tradeoff faced by designers of visual displays. Although humans can easily discern spatial details of even less than 1' of visual angle (corresponding to a Snellen visual acuity of 20/20 or better), most virtual environment displays cannot provide such fine spatial resolution.

Some potential solutions to the FOV versus spatial resolution tradeoff problem include: (a) tiling multiple displays (e.g., the multiple LCS of Kaiser Electro-optics HMDs); (b) using a mixed spatial resolution to provide a higher spatial resolution within the central 2 degrees to 5 degrees of the visual field (where human visual acuity is best) but lower spatial resolution outside this region (e.g., CAE Electronics' high-priced fiber-optic HMD); (c) use binocular or biocular displays with partial overlap of the visual fields instead of complete overlap; (d) using multiple projections onto surrounding walls or screens (e.g., the CAVE or the Virtual Workbench); or (e) designing a new and better display element.

Each of these potential solutions has additional technological problems and each can significantly increase the complexity and cost of the display. For example, a mixed resolution display assumes *line of sight* viewing -- where the user *always* looks straight ahead rather than off the to side. Moreover, the sense of immersion created by a wider field of view does not necessarily result in better performance or greater user comfort.

Stereoscopic displays vs. biocular displays. Others feel that the 3D stereoscopic imagery created with binocular displays can provide a compelling sense of presence and that honing stereoscopic displays is the most pressing VE challenge. Stereopsis certainly can enhance performance of close work involving fine motor control, such as surgical operations, as well as long-range distance perception of an object one mile away from another at optical infinity. In a binocular display a slightly different image is presented to the right eye from that presented to the left

eye, unlike a biocular display in which exactly the same image is presented to both eyes. Stereopsis results from the fusion of the two slightly different views that the user's laterally-displaced eyes receive in viewing a real or a simulated 3D stereoscopic display (Arditi, 1986; Davis & Hodges, 1995; Tyler, 1983).

Because stereoscopic depth depends on the interpupillary distance (IPD) between the two eyes, the optics of the binocular display or the binocular parallax of the computer graphics and/or the offset of the two visual sensor devices should be adjustable. The average user's IPD is approximately 63 mm, with a range of 53 mm to 73 mm (Kalawsky, 1993). The larger the IPD, the more salient the perceived stereoscopic depth -- so artificially increasing the *virtual* IPD could enhance the perceived stereoscopic depth. But, beware! Creating too much binocular parallax can destroy perceptual fusion, resulting in diplopia or double-vision with accompanying feelings of nausea, eye strain, and possible performance deficits (Piantanida, Boman, & Gille, 1995).

Using a binocular (or biocular) visual display with only partial overlap can result in a wider FOV, while maintaining an adequate spatial resolution, as previously explained. For most humans, in viewing the real world, the region of binocular overlap is 120°, with a monocular visual field of approximately 35° flanking each side of the binocular overlap. For most VE systems with partial overlap, however, the region of overlap is considerably less than 120°. Also, with partial overlap there are flanking right and left monocular regions of the display; each monocular region provides very different spatial and luminance configurations for the right versus the left eye. Because of these two factors, the user perceives binocular rivalry and monocular suppression within each flanking monocular region of the display, with concomitant discomfort and disorientation. Potential solutions to these problems include (1) a relatively large region of overlap, one that more closely approximates the 120° binocular field of natural viewing; (2) blurring the edges of the display so that the entire visual field appears less fragmented (Melzer & Moffitt, 1989; Melzer & Moffitt, 1991); and (3) using *convergent* rather than *divergent* binocular displays (Melzer & Moffitt, 1989; Melzer & Moffitt, 1991). As a *rule of thumb* Kalawsky (1993) recommends using complete binocular overlap when moderate FOVs suffice, but partial overlap only if very wide FOVs are necessary.

Binocular displays are heavier, more complex, and costlier than are simpler biocular displays, although both require adjustable IPDs and careful alignment of each eye's view. Still, not everyone possesses good stereopsis so binocular displays

have limited effectiveness for these individuals. There are amblyopes who have no stereopsis as well as some individuals who are stereoanomalous. Richards (1971) has shown that stereoanomalous observers may be able to perceive stereoscopic depth for objects presented in front of the point of fixation, but not behind it (or vice versa). Even some individuals with normal stereopsis may have difficulty perceiving stereoscopic depth in a visual display, although with training and feedback their stereoscopic visual performance sometimes improves dramatically (McWhorter, 1993; Surdick, Davis, King, & Hodges, 1997).

So, given the above constraints, when are binocular VEs more useful than biocular ones? Binocular VEs are useful when a visual scene is presented in a perspective view than in a bird's eye view (Barfield & Rosenberg, 1992; Yeh & Silverstein, 1992), when monocular cues provided by a biocular display provide ambiguous or less effective information than the stereopsis provided by binocular HMDs (Yeh & Silverstein, 1992), when static or slowly-changing visual displays are used rather than rapidly-changing, dynamic displays (Wickens, 1990; Wickens & Todd, 1990; Yeh & Silverstein, 1992), when ambiguous objects or complex scenes are presented (Cole, Merritt, & Lester, 1990; Spain & Holzhausen, 1991), and when complex 3D manipulation tasks require ballistic movements or very accurate placement and positioning of objects or tools. Stereopsis is helpful in these situations for a two primary reasons. First, it helps disambiguate elevation and distance information in providing information about the spatial layout of objects (e.g., in a perspective view). Second, it provides the user with fine depth discrimination for objects (and the shapes of objects) located within an arm's length of the user.

Color versus monochrome. The issue of color displays in virtual environments is whether one should use color or not. The use of color within the display can add to a sense of immersion and realism. Users often prefer color displays and color-coding within a display can be helpful, especially in virtual environments where both the rate and amount of information transmitted is high (Christ, 1975; Stokes, et al., 1990). In short, some reasons color displays are desirable include: (1) color can be used to unify or cluster disparate elements of a display (Christ, 1975; Christ & Corso, 1982); (2) color seems to be processed earlier and faster than other types of visual information, such as shape; (3) objects are more easily identified by color than by other features based on size, shape, or brightness; (4) color coding can significantly reduce visual search time; and (5) color adds to the *realism* of the virtual environment.

Given all of these advantages, why would one consider *not* using a color display within a virtual environment? The reasons not to use color primarily involve technological limitations and cost factors. With most existing displays the use of color involves tradeoffs with other capabilities of the display such as spatial resolution or temporal resolution. Available color displays lack the full range of chromaticities (hues and saturations) that the human can perceive. They also tend to have less contrast or brightness than monochromatic or grayscale displays as well as being heavier and costlier. Although approximately 8% of the male adult population have inherited color vision deficiencies that limit the usefulness of color displays[2], most adults (both male and female) will *acquire* color vision deficiencies as a result of normal aging processes or of disease processes (e.g., diabetes). Moreover, in low luminance environments, such as those simulating nighttime vision, all users are blind to color -- instead they see only in black and white and shades of gray.

Temporal resolution, update rates, and time lags. Still other VE developers and users feel that time lags caused by relatively slow update rates or frame rates are the most pressing challenge. Both frame rate and update rate affect the temporal resolution of a visual display. Frame rate is a hardware-controlled variable that determines how many images each eye sees per second (measured in Hertz), whereas update rate is the rate at which changes in the image are updated. Of the two, update rate is the more problematic. The update rate depends on the computational complexity of an image[3] and can be no faster than the frame rate. The update rate constrains how fast virtual objects can move around in virtual space yet avoid jerky movement.[4] Update rate also constrains how quickly the user can move his or her head yet have the virtual images remain meshed with the head movement.[5] If the temporal resolution of the visual display is too low, it can hinder task performance or cause illusory motion artifacts. In fact, noticeable time lags in update rates can cause nausea and disorientation in space -- making the user

---

[2]Only about 0.5% of females have inherited color deficiencies.

[3] Update rates also are influenced by time lags and delays in sensory devices (e.g., headtracking devices). The effects of computational complexity of images and of sensory delays are additive.

[4] The minimum update rate in Hertz is the virtual object's angular speed in arcmin/sec divided by 15, assuming a maximum displacement of 15' per frame for the perception of smooth movement (e.g., Braddick, 1974).

[5]When the user's head position is monitored and the visual scene is updated according to the head position, then update rate depends on sensor lag in determining the head position as well as on the computational complexity of the visual imagery. These two factors are additive.

very uncomfortable and, perhaps, hindering his or her performance (Piantanida, et al., 1995).

Meshing real and virtual imagery. In augmented-reality all or part of a real-world scene is combined with synthetic imagery. There are two approaches to creating this augmented reality. The most popular augmented reality system is to use a see-through head-mounted display in which computer graphics or symbology is superimposed on a direct view of the real world. An optical combiner composed of half-silvered mirrors is used to mesh the synthetic imagery with real-world scenes (Barfield, Rosenberg, Han, & Furness, 1992; Davis, 1996). Because the older heads-up display (HUD) is an example of this type of augmented reality, alot already learned about presenting symbology in a HUD is relevant here.

Another augmented reality system is to use an opaque HMD or projection screen system in which computer graphics is superimposed on a local or remote video of real-world scenes. Video keying is used to electronically merge video and synthetic images, similar to that used for CNN's news broadcasts. Computer graphics or other video clips are used to fill-in a large blue screen behind the anchor. The particular blue hue is used to mark the area for video keying because most human skin tones do not contain this color (Barfield, Rosenberg, & Lotens, 1995). (However, if the anchor wore a dress or jacket of that same blue color, TV viewers would have an eerie view of a vanishing anchorperson!)

Both methods of visually presenting augmented reality have some similar requirements. For both, accurate registration of the synthetic imagery and the real world is necessary. For example, meshing of perceived distance and of spatial layout matters. Roscoe (1984; 1993) has reported that virtual objects often appear farther away than do real-world objects; he suggests using a magnification factor of approximately 1.25 to correct for this disparity in the perceived distances of virtual and real objects. Although some have confirmed Roscoe's findings in an augmented reality system (Rolland, Gibson, & Ariely, 1995; Rolland, Holloway, & Fuchs, 1994), others have reported that in see-through HMDs virtual objects almost always appear closer (e.g., K. Moffitt, personal communication). Perhaps virtual objects appear closer in a see-through HMD because virtual objects occlude real-world objects or because accommodation cues signal that the virtual object presented on a visor is closer than the real objects viewed through the visor. In fact, when spatial layouts and updates of virtual imagery does not mesh with that of real-world scenes, this may introduce a new source of *simulator sickness* which includes vertigo, dizziness, and disorientation. It is also true that directly viewed real-world

scenes in a see-through HMD will mesh with vestibular cues resulting from turning one's head, et cetera; thus augmented-reality displays with see-through HMDs may reduce some of the simulator sickness symptoms reported with virtual environments created with opaque HMDs and updates rates that are "not fast enough".

For see-through HMDs there are some additional problems because the real-world scene effectively provides a background against which the synthesized images and symbology are viewed. At nighttime and under dim illumination, the effective background luminance provided by the real environment can be quite low, increasing the effective contrast and visibility of light symbols, but decreasing those of dark symbols. On a bright, sunny day, however, the effective background luminance provided by the real environment can be quite high, drastically decreasing the contrast and visibility of even the brightest synthetic symbology. To offset this problem one could change the overall gain of the real-world luminances to some constant value and, perhaps, simultaneously change the mean luminance and contrast gains in the synthetic imagery. Some lessons learned from using heads-up displays (HUDs) are relevant here -- the military currently uses a dark visor over the headset combiner to overcome the problem of a bright, ambient luminance from the real-world scene. Another possible approach would be to use light-sensitive filters over the headset combiner -- similar to that used in modern, high-tech sunglasses. Yet another possibility is the ability to change the contrast and polarity (i.e., light versus dark) of the virtual symbology.

We have emphasized visual augmented reality systems, which are the most highly developed augmented reality systems to date. Of course, augmented reality systems using other sensory modalities, such as auditory and tactile, are possibilities and also are worth exploring. After all, virtual environments are a multisensory and interactive experience. In the everyday real world individuals are rarely confronted with only visual information, but instead are simultaneously bombarded with a wealth of information from a variety of sensory modalities, including auditory information.

B. Some auditory considerations

While perception of 3D objects in VEs can be created using a wide variety of visual information, auditory inputs may be synchronized with visual inputs to provide an even greater sense of immersion and realism than that achieved by a single sensory modality (Bryson, Pausch, Robinett, &

van Dam, 1993; Burgess & Verlinden, 1993; Kalawsky, 1993). Although VE development has focused more on visual aspects of the environment than on other potential components of the system, with the introduction of more sophisticated and cost-effective audio hardware and system software it is now possible and prudent to use auditory information to enhance the realism and sense of presence within virtual worlds (Astheimer, 1993; Barfield & Furness, 1995; Bargar, Blattner, Kramer, Smith, & Wenzel, 1993).

Spatialized vs. non-spatialized (stereo) sound. Whether to include 3D spatial audio or simple stereo sound in simulated environments is an important issue for incorporating auditory displays into VE systems. Whereas simple stereo sound elicits a perceived position of the sound source along a one-dimensional line (i.e., the sound is heard as coming from the left or the right of the listener), spatialized sound is perceived as coming from a precise location in space and has not only a left-right attribute but also up-down and distance components (Barfield & Furness, 1995; Bryson, et al., 1993; Burgess, 1992).

Spatialized or true 3D sound was first created using a digital-signal processing device known as the Convolvotron, which was designed by Crystal River Engineering and NASA Ames Research Center (Bryson, et al., 1993; Burgess, 1992). This commercialized system transforms digital auditory signals in a similar manner to how the shape of the human's outer ear (or pinna) transforms real-world sounds (Barfield & Furness, 1995). That is, the Convolvotron uses head-related transfer functions (HRTFs) to filter the digitized sounds (Astheimer, 1993; Burgess, 1992). These HRTFs not only take into account pinnae effects but also interaural time differences (ITDs), interaural intensity differences (IIDs), and diffraction effects caused by the presence of the listener's head and body (Barfield & Furness, 1995; Durlach, 1991; Kalawsky, 1993). When the filtered sounds are heard by a listener, they appear to have a true sense of 3D location in space. Not only do these sounds have a directional component, but they also retain the acoustical properties of the listening environment (e.g., room reverberation cues) (Bryson, et al., 1993).

Spatialized sound is capable of providing an incredible sense of realism and presence in a VE. Furthermore, because of its spatial characteristics, it can be used to assign qualities to virtual objects. Spatial 3D audio is also very useful because it can provide crucial directional cues that better allow us to

navigate within the virtual world. Another advantage that spatialized sound has over simple stereo sound is that when the former is presented over headphones, for example within a HMD, it is perceived as coming from a particular location in space outside of the listener's head (Burgess, 1992). Simple stereo sound, however, is often perceived as originating from within the user's head (Barfield & Furness, 1995; Burgess, 1992). This perceptual problem, known as lack of externalization, is the direct result of stereo recording's poor model of how the human's pinnae alter sound stimuli once it arrives at the ear (Burgess, 1992).

Although spatialized sound can greatly enhance the spatial characteristics of a virtual auditory environment, it does have its drawbacks. In order to create spatialized sound, the acoustical properties of a particular listening environment must be modeled so that reverberation cues can be provided (Burgess, 1992; Burgess & Verlinden, 1993). Once a model has been developed, it must then be updated in real time. An even bigger drawback for spatialized sound is that with the Convolvotron's fast and complex computing power also comes its high cost. Incredible hardware costs have made this special processing device impractical for many VE system designers (Burgess, 1992). Despite the fact that some designers may choose to invest in a lower cost system, such as the presently-available Focal Point system, many system designers still choose to use simple stereo sound. Better, more affordable audio hardware and software is necessary before spatialized sound will be widely used by the VE community.

Real-world vs. synthesized sound effects. Must sounds used in VEs resemble ordinary real-world sounds or can artificially-created, synthetic sounds relay the same intuitive information in virtual worlds? The purpose of simulated environments, such as VEs, is not to create worlds which totally emulate the real world; however, system developers should be able to design realistic worlds if that is indeed the goal (Bargar, et al., 1993). Therefore, we probably want the capability of including both real-world and synthesized sound effects, depending upon the specific task employed within the VE and the precise goal of the environmental design.

Studies have revealed that users are more likely to respond favorably to real-world sounds than to artificial ones (Bargar, et al., 1993). In fact, natural sounds are often considered an effective type of auditory input in simulated environments merely because they are immediately recognizable by the

listener and therefore require neither adaptation nor training (Durlach, 1991). Although synthesized sounds may not be familiar to the listener, they can provide a greater range of distinct sound combinations than can real-world sounds (Bryson, et al., 1993). However, presenting virtual sounds that relate to the listener's everyday world experiences and that can be used in intuitive, self-evident ways is probably the most efficient and effective use of auditory input in simulated auditory display systems.

Meshing real and virtual auditory input. As previously mentioned, augmented reality displays do not have to be specifically visual in nature. In fact, auditory augmented reality systems can be used to enhance the real world in novel ways that allow the human to more efficiently and effectively interact with complex systems (Barfield & Furness, 1995). For example, digitized spatial audio can be overlaid onto a real sound in order to amplify the significance of the real-world noise. Conversely, anti-noise filters can dampen predictable or regular real-world noise.

C. Meshing multisensory inputs and Interactive Displays

Just as real and virtual imagery within a given sensory modality must mesh, so too must the different sensory inputs mesh together so that a coherent percept of the virtual world results. This is especially true for interactive displays in which the user moves around within the virtual environment and manipulates virtual objects. Time lags between inputs to the different sensory modalities can be confusing, disorienting, and downright nauseating. Not only must the temporal properties mesh across the different sensory modalities, but also the spatial properties must properly mesh together. Accurately perceived locations of objects in the azimuth, elevation, and distance (x, y, z) dimensions are important for the perception of where the objects are located and of the spatial layout of scenes or configurations. What happens when information provided by the different sensory modalities conflicts or is not in register across the senses? If visual and auditory information do not mesh together to form a coherent percept, often the visual sense will dominate -- a phenomenon known as visual capture (Welch & Warren, 1986). Although vision is our keenest sense, when it conflicts with tactile and kinesthetic inputs, it is the latter which may dominate (Welch & Warren, 1986).

Exactly how information is integrated within a given sensory modality or across different sensory modalities remains an unsolved puzzle. The solution to this puzzle can help to determine the relative weight of each sensory input in forming a coherent percept of the real or virtual world. It also can help us

determine how much error the system can tolerate in meshing together various sensory inputs.

Interacting with the environment sometimes can compensate for poor sensory or perceptual information within a virtual environment. For example, a study by Smets and Overbreke (1995) found that yoking a camera's view of a visual scene to the user's head movements improved the user's ability to solve a visual puzzle, despite relatively poor spatial resolution in the visual image. Results such as these suggest an ecological or Gibsonnian (Gibson, 1979) approach to the development and use of virtual environments: Performance within a virtual environment partially depends on the interactions between the perceptual quality of the displays and active manipulation of the environment.

In general, the more capabilities the virtual environment has, the more difficult it is to effectively interface this system with the human user. Yet, the more capabilities the virtual environment has, especially if these capabilities are calibrated with each other and mesh together, the more compelling the percepts formed of the objects, spatial layouts, scenes and configurations in the virtual environment.

References

Arditi, A. (1986). Binocular vision. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of Perception and Human Performance, Vol. 1 New York: Wiley.

Astheimer, P. (1993). What you see is what you hear: Acoustics applied in virtual worlds. In Proceedings of the IEEE 1993 Conference.

Barfield, W., & Furness, T. A. (Eds.). (1995). Virtual Environments and Advanced Interface Design. New York: Oxford University Press.

Barfield, W., & Rosenberg, C. (1992). Comparison of stereoscopic and perspective display formats for exocentric judgment tasks. In Paper presented at the Annual Meeting of the Human Factors Society. Atlanta, Georgia.

Barfield, W., Rosenberg, C., Han, S.-H., & Furness, T. (1992). A god's eye (exocentric) versus a pilot's eye (egocentric) frame-of-reference for enhanced siuational awareness. In Proceedings of the Annual Meeting of the Human Factors Society.

Atlanta, Georgia: Working Paper: Department of Industrial Engineering, University of Washington.

Barfield, W., Rosenberg, C., & Lotens, W. A. (1995). Augmented-reality displays. In W. Barfield & T. A. Furness (Eds.), Virtual Environments and Advanced Interface Design New York: Oxford University Press.

Bargar, R., Blattner, M., Kramer, G., Smith, J., & Wenzel, E. (1993). Effective nonspeech audio for virtual reality. In Proceedings of the IEEE 1993 Conference.

Bryson, S., Pausch, R., Robinett, W., & van Dam, A. (1993). Implementing virtual reality. In Proceedings of the ACM SIGGRAPH 1993 Conference. Anaheim, CA.

Burgess, D. A. (1992). Techniques for low cost spatial audio. In Proceedings of the Fifth Annual Symposium on User Interface Software and Technology, . New York, NY: ACM.

Burgess, D. A., & Verlinden, J. C. (1993). An architecture for spatial audio server No. GIT-GVU-93-91). Graphics, Visualization, and Usability Center, Georgia Institute of Technology.

Christ, R. E. (1975). Review and analysis of color coding research for visual displays. Human Factors, 17, 542-570.

Christ, R. E., & Corso, G. (1982). The effects of extended practice on the evaluation of visual display code. Human Factors, 25, 71-84.

Cole, R. E., Merritt, J. O., & Lester, P. (1990). Remote manipulator tasks impossible without stereo TV. In J. O. Merritt & S. S. Fisher (Eds.), Proceedings of the SPIE: Stereoscopic Displays and Applications I, 1256 (pp. 255-265). Bellingham, WA: International Society for Optical Engineering.

Davis, E. T. (1996). Visual requirements in HMDs: What can we see and what do we need to see? In J. Melzer & K. Moffitt (Eds.), Head Mounted Displays: Designing VR System for the User New York: McGraw-Hill.

Davis, E. T., & Hodges, L. F. (1995). Human stereopsis, fusion, and stereoscopic virtual environments. In W. Barfield & T. Furness (Eds.), <u>Virtual Environments and Advanced Interface Design</u> New York: Oxford.

Durlach, N. (1991). Auditory localization in teleoperator and virtual environment systems: Ideas, issues, and problems. <u>Perception, 20,</u> 543-554.

Gibson, J. J. (1979). <u>The ecological approach to visual perception</u>. Boston: Houghton Mifflin.

Kalawsky, R. S. (1993). <u>The Science of Virtual Reality and Virtual Environments</u>. Wokingham, England: Addison-Wesley Publishing Company.

Melzer, J. E., & Moffitt, K. (1989). Partial binocular-overlap in helmet-mounted displays. In <u>SPIE Proceedings: Display System Optics II,</u> 1117 (pp. 56-62). San José, California: SPIE -- The International Society for Optical Engineering.

Melzer, J. E., & Moffitt, K. (1991). An ecological approach to partial binocular overlap. In <u>S.P.I.E. Proceedings: Large Screen Projection, Avionics, and Helmet-mounted Displays,</u> 1456 (pp. 175-191). San José, California: SPIE -- The International Society for Optical Engineering.

Norman, D., & Bobrow, D. (1975). On data-limited and resource-limited processing. <u>Journal of Cognitive Psychology, 7,</u> 44-60.

Piantanida, T., Boman, D. K., & Gille, J. (1995). Perceptual issues in virtual reality. Unpublished manuscript.

Richards, W. (1971). Anomalous stereoscopic depth perception. <u>Journal of the Optical Society of America, 61</u>(3), 410-414.

Rolland, J. P., Gibson, W., & Ariely, D. (1995). Towards quantifying depth and size perception in virtual environments. <u>Presence, 4</u>(1), 24-49.

Rolland, J. P., Holloway, R. L., & Fuchs, H. (1994). A comparison of optical and video see-through head-mounted displays. In <u>S.P.I.E.,</u> 2351 (pp. 293-307).

Roscoe, S. N. (1984). Judgments of size and distance in imaging displays. <u>Human Factors,</u> <u>26</u>(6), 617-629.

Roscoe, S. N. (1993). The eyes prefer real images. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.), <u>Pictorial Communication Virtual and Real Environments</u> Washington, D.C.: Taylor & Francis.

Smets, G. J. F., & Overbeeke, K. J. (1995). Trade-off between resolution and interactivity in spatial task performance. <u>IEEE Computer Graphics and Applications,</u> <u>September,</u> 46-51.

Spain, E. H., & Holzhausen, K. P. (1991). Stereoscopic versus orthogonal view displays for performance of a remote manipulation task. In J. O. Merritt & S. S. Fisher (Eds.), <u>Proceedings of the SPIE: Stereoscopic Displays and Applications II,</u> 1457 (pp. 103-110). Bellingham, WA: International Society for Optical Engineering.

Stokes, A., Wickens, C., & Kite, K. (1990). <u>Display Technology: Human Factors Concepts.</u> Warrendale, PA: Society of Automotive Engineers, Inc.

Surdick, R. T., Davis, E. T., King, R. A., & Hodges, L. F. (1997). The perception of distance in simulated displays -- A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. <u>Presence.</u>

Tyler, C. W. (1983). Sensory processing of binocular disparity. In C. M. Schor & K. J. Ciuffreda (Eds.), <u>Vergence Eye Movements: Basic and Clinical Aspects</u> Boston: Buttersworth.

Welch, R. B., & Warren, D. H. (1986). Intersensory interactions. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), <u>Handbook of Perception and Human Performance</u> New York: Wiley.

Wickens, C. D. (1990). Three-dimensional stereoscopic display implemention: Guidelines derived from human visual capabilities. In J. O. Merritt & S. S. Fisher (Eds.), <u>Proceedings of the SPIE: Stereoscopic Displays and Applications I,</u> 1256 (pp. 2-11). Bellingham, WA: International Society for Optical Engineering.

Wickens, C. D., & Todd, S. (1990). Three-dimensional display technology for aerospace and visualization. In <u>Proceedings of the Human Factors Society 34th Annual Meeting,</u> (pp. 1479-1483). Santa Monica, CA: Human Factors Society.

Yeh, Y. Y., & Silverstein, L. D. (1992). Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. <u>Human Factors,</u> <u>34</u>(5), 583-600.

Keith Eckerman's report was not available at the time of publication.

AVIATION FUEL IDENTIFICATION

PART I.
NEURAL NETWORK ANALYSIS
OF THE
CONCENTRATION OF BENZENE AND NAPHTHALENE DERIVATIVES
IN FUELS

PART II.
THE RESPONSE OF AN ARRAY OF
VAPOR SENSITIVE DETECTORS
TO
WEATHERED FUELS

Paul A. Edwards
Associate Professor
Department of Chemistry

Edinboro University of Pennsylvania
Cooper Hall
Edinboro, PA 16444

AVIATION FUEL IDENTIFICATION

PART I.
NEURAL NETWORK ANALYSIS
OF THE
CONCENTRATION OF BENZENE AND NAPHTHALENE DERIVATIVES
IN FUELS

PART II.
THE RESPONSE OF AN ARRAY OF
VAPOR SENSITIVE DETECTORS
TO
WEATHERED FUELS

Paul A. Edwards
Associate Professor
Department of Chemistry
Edinboro University of Pennsylvania

Abstract

Part I

The concentration of ten benzene and naphthalene derivatives was determined in one hundred and two (102) aviation fuels using gas chromatography/mass spectrometry (GC/MS). Artificial neural network analysis techniques were then used to classify the fuels from the resultant data matrix. Ultimately only six compounds were required to make the classification.

Part II

A preliminary study was done comparing the response of an array of vapor sensitive detectors to *fresh* and *weathered* JP-4, JP-5, JP-8, JPTS and JetA aviation fuels. The fuels were *weathered* by allowing them to evaporate in a hood for two weeks. The response of the array was essentially the same for *weathered* and *fresh* samples of the same fuel. These results must be confirmed in a more complete study. However, if this is correct, the response of such an array could be used to identify samples of aviation fuels **almost independent of the age of the sample.**

Part I.

Introduction

The development of rapid and reliable methods for identifying aviation fuels in environmental matrices continues to be a primary goal of research at the Environics Directorate of Armstrong Laboratory. This group and other researchers report identifying fuels based on their gas chromatography (GC) *fingerprint* [1] and their composition as determined by gas chromatography /mass spectrometry (GC/MS) techniques [2-5]. Efforts continue to improve the accuracy of and reduce the time required for these analyses, because the results have legal and financial implications for polluters and those involved in cleanup.

Aviation fuels are complex mixtures of three broad categories of petroleum hydrocarbons: alkanes, alkenes, and aromatic hydrocarbons. Using GC/MS methods it is possible to move beyond these broad categories, and identify and quantify a large number of the chemical compounds present. Using mathematical methods known as *pattern recognition*, it is possible to use the composition of the fuels as a means of identification. A critical goal of this particular research was to be able to identify fresh and *weathered* fuels. Thus, the focus was on compounds that were likely to persist in a sample and **not** likely to be affected by *weathering*.

There are three general modes by which an aviation fuel can *weather*: evaporation, dissolution in water, and degradation. Thus the compounds to consider should be insoluble in water, non-volatile, and not readily oxidizable by air or biological activity. Obviously the compounds should also be known to be present in the majority of aviation fuels in experimentally reasonable amounts. Polyaromatic hydrocarbons such as naphthalene and alkyl-substituted naphthalene derivatives were considered to be satisfy these requirements.

Artificial neural networks (ANNs) are computer simulations of biological nervous systems [6]. A schematic representation of an ANN is displayed in Figure 1. In general terms, numerical information enters an ANN through a layer of input neurons or nodes and exits through a layer of output nodes. Information passes from the input to the output layer through a hidden layer (or layers). As information is passed through the layers, numerical weights, biases and transfer function are applied that adjust the

connections between the nodes such that the information exiting the network has the desired value or pattern.

Figure 1: Schematic Representation of an Articial Neural Network
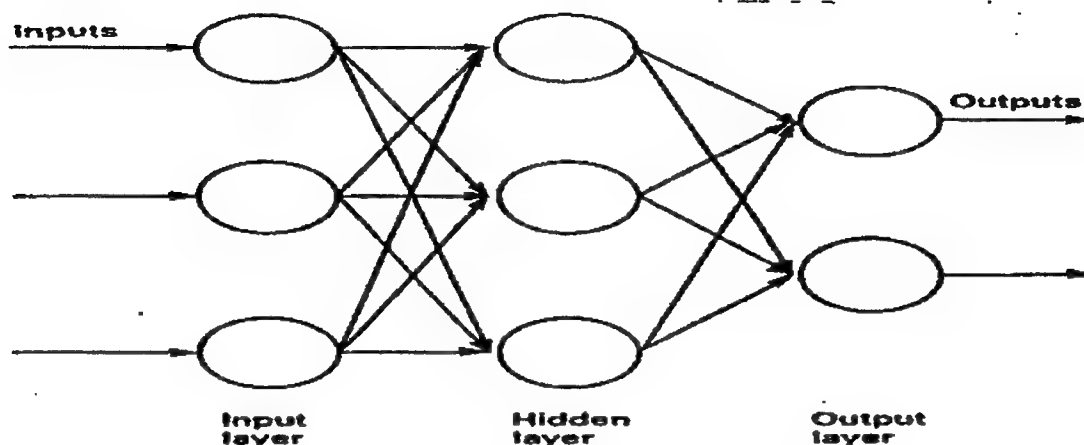
Artificial neural networks have already been applied to the classification of fuels based on chromatography data [2,3], laser induced fluorescence spectra [7], and the response of vapor sensitive metal oxide sensors [8]. The application of ANNs to classification problems such as this involves at least three steps: training, architecture optimization, and validation. Training is the automated process of adjusting the weights and biases in the network so that the output patterns generated by the network match those in a training data set. Back-propagation is one of the more commonly used computational formalisms for training and Wythoff has written an excellent tutorial on these methods [9]. Training is interconnected with architecture optimization, which is the process of determining the number of input, hidden and output nodes required in the ANN to yield the best classification results. Training and architecture optimization continue until an architecture is found that properly classifies the samples in the training set, and, at least in this laboratory, has a minimum size. Validation is the process of demonstrating the ability of the trained network to correctly classify patterns not used in the training process. A number of commercial and even *shareware* software packages are available for accomplishing these tasks including, FIP, a "Fuel Identification Program" by Faruque *et al* [3], the package developed and used in this laboratory.

Experimental Section

The Data Set: A set of one hundred and five (105) aviation fuels was collected from Wright Patterson AFB Energy Management Laboratory, Mukilteo, WA, Energy Management Laboratory, MacDill AFB, and Wright Patterson Aerospace Fuel Laboratory. The classes of fuels represented in the data set, and the number of samples in each class, are displayed in Table 1.

Table 1: The Aviation Fuel Data Set

No. of Samples	Category
20	JP-4
13	JP-5
8	JP-7
20	JP-8
11	JPTS
20	Jet A
13	AvGas

Sample Preparation: Samples of *fresh* fuels were prepared by diluting 100 μL of fuel with 100 μL of an internal standard and 800 μL of methylene chloride. The internal standard solution consisted of 25 mg of d_{10}-anthracene and 25 μL of d_{10}-ethylbenzene.

A set of twenty-six vials (two each of thirteen fuels) was *weathered* by allowing them to evaporate in a hood. One vial was used to track the mass of the fuel as it *weathered*. Aliquots (100 μL) of the *weathered* fuel were removed from the second and diluted for GC/MS analysis as stated above.

Gas Chromatography/Mass Spectrometry (GC/MS): The samples were analyzed by GC/MS, utilizing an instrument calibration method designed to provide accurate quantitative measurements of the concentration of the selected benzene and naphthalene derivatives in fuel samples. Quantitation was based on the d_{10}-anthracene, and the d_{10}-ethylbenzene provided a quality control measurement, being used along with the d_{10}-anthracene to reveal quickly analysis runs that had failed due to instrument accident or artifact.

The samples were analyzed using a benchtop gas chromatograph (GC) /mass selective detector (MSD) system (HP-5890 GC and HP-5970 MSD). The GC and MSD were interfaced by a direct capillary inlet. The gas chromatograph was equipped with a split/splitless injection port for use with capillary

columns and modified to an HP-5890 series II configuration with electronic pressure control. A personal computer (HP Vectra 486/66 MHz) running HP DOS/Windows MS-Chemstation Software was used for system control and data acquisition. Injections of prepared samples were made with an HP-7673A autosampler.

Gas chromatographic separations were made using a fused silica capillary column, 20m long, with an internal diameter of 0.1mm, and coated with 0.4μm of bonded and crosslinked polymethylsiloxane with 5%-phenyl substitution (DB-5, J&W Scientific, Inc.). Helium carrier gas was supplied with a constant head pressure of 45 psig. The temperature program conditions are summarized in Table 2. The capillary column was plumbed directly into the ion source of the mass selective detector using a direct capillary interface. Injections were made with the injection port operating in the split mode.

Table 2: Gas Chromatography Conditions

Oven Program Rate:	15°C/min.
Final Oven Temperature:	270°C
Final Oven Hold:	5 min.
Injection Port Temp:	300°C
MS Interface Temp:	225°C
Initial Oven Hold:	4 min.
Carrier Gass:	He
Head Pressure:	45 psig
Injection Volume:	1μL
Split Port Flow:	60 mL/min.

The MSD was a quadrupole mass spectrometer with an electron impact ion source. It was scanned from 35 amu to 500 amu using mass spectrometry conditions summarized in Table 3. Selected analytes were quantified using the EnviroQuant target analyte software provided by the Chemstation computer, and using target ions and retention times established for the selected analytes listed in Table 4 by the analysis of standard mixtures.

Table 3: Mass Spectrometry Conditions

Solvent Delay:	4 min.
Electron Multiplier:	1800 volts
Ionization Method:	Electron Impact
Ionization Voltage:	70 volts
Low Mass:	35 amu
High Mass:	500 amu
A/D Duplicate Samples:	4 samples/axis point
Scanning Rate:	1.3 scans/sec.

The concentrations of each of the selected benzene and naphthalene derivatives in a sample were stored as a $1 \times n$ data vector representing a concentration profile or *pattern* for that sample. The resultant $105 \times n$ data matrix was then analyzed using artificial neural network methods.

Table 4: Analytes Identified and Quantitated

Compound	Quantitation Standards #1	Quantitation Standards #2
naphthalene	x	
1-methylnaphthalene	x	
2-methylnaphthalene	x	
1,2-dimethylnaphthalene	x	
1,3-dimethylnaphthalene	x	
1,4-dimethylnaphthalene	x	
1,5-dimethylnaphthalene	x	
1,8-dimethylnaphthalene	x	
2,3-dimethylnaphthalene	x	x
2,6-dimethylnaphthalene	x	x
1-ethylnaphthalene	x	x
2-ethylnaphthalene	x	
1,4,6,7-tetramethylnaphthalene	x	x --
1,2,3,4-tetrahydronaphthalene	x	
benzene		x
isopropylbenzene		x
1,2-diethylbenzene		x
1,2,3,4-tetramethlybenzene		x

Neural Network Analysis: Neural network analysis was accomplished using FIP, a "Fuel Identification Program" by Faruque *et al* [3] on a Sun Microsystems SPARCstation2 super minicomputer operating under SunOS 4.1.1. FIP is a suite of MATLAB[tm] [10] and Neural Network TOOLBOX[tm][6] functions package developed and used in this laboratory for classification of fuels based on gas chromatography data. Back-propagation (BP) methods were used to train the neural network, unless otherwise indicated. As implemented in FIP, the parameters controlled by the use include the architecture of the network, the

number of iterations in each cycle in the calculation, the momentum, the learning rate, and the maximum and minimum error goals. The number of iterations in each cycle was generally set to 1000. The momentum was 0.95 and the learning rate was 0.1, Initial values for the weights and biases were randomly generated at the start of each training cycle and adjusted during each iteration. If the number of classifications after 1000 iterations was greater than the maximum error goal, the values of the weights and biases were reinitialized, and training continued for another cycle of 1000 iterations. Training terminated when the number of misclassifications was between the minimum and maximum error goals set for that calculation.

Results and Discussion

One sample was removed from consideration after examining its GC profile, composition and *weathering* behavior. The fuel, identified by its supplier as a JP-4, behaved in all respects as a JP-8.

It was not possible to classify all the fuels by artifical neural network (ANN) analysis using a single-step approach to the problem. Attempts were made with input patterns containing the concentrations of naphthalene and its derivatives, and the concentrations of a combination of benzene and naphthalene derivatives. Attempts were also made with input patterns containing the ratios of the concentration of naphthalene derivatives relative to that of naphthalene, and with descriptors formed by principal component analysis of the raw data. The network architecture used in these attempts included as few as six and as many as twelve input nodes of information. The number of nodes in the hidden layer was also varied from as few as six to as many as eight. Training did not proceed to fewer than two(??) misclassifications using the full data matrix.

The main block to training appears to be the similarity between JP-8 and JetA fuels. JetA and JetA-1 are kerosene-based commercial turbine fuels. The major difference in properties is that JetA-1 has a slightly higher maximum freezing point (-47°C) than JetA (-50°C) [11]. JetA is the commercial turbine fuel used in the United State while JetA-1 is used primarily outside the U.S. JP-8 is a kerosene-based military turbine that is similar to JetA-1. JetA-1 is also designated as F-35 in NATO fuel specifications.

Given the similarity of JetA, JetA-1, and JP-8, the overlap of these groups noted in the display of the first three principal components of the data set becomes understandable.

Training was then attempted with two different two-step approaches to the problem. One two-step approach involved first dividing the data set into three groups: *special* fuels, *kerosene* fuels, and aviation gasoline. The *special* fuels group included fuels being phased out of routine usage or used by only a limited number of aircraft, *ie.* JP-4, JP-7, and JPTS. The *kerosene* fuels group included turbine fuels in common usage, *ie.* JP-5, JP-8 and Jet-A. The aviation gasoline group contained only 100/130 octane fuel (AvGas) used by piston aircraft. The second step in this approach involved divided the *special* and *kerosene* fuel groups into their respective classes. The second two-step approach involved first dividing the fuels into the classes JP-4, JP-5, JP-7, JPTS, AvGas and a combined JetA/JP-8 class. The second step in this approach involved divided the combined JetA/JP-8 class into the respective two classes.

The approach involving first classifying the fuels as *special* fuels, *kerosene* fuels or aviation gasoline was partially successful. Networks were trained using the concentrations of naphthalene and naphthalene derivatives as input patterns that could 100% correctly classify fuels into the broad categories and fuels in the *special* category. However, classifying fuels in the *kerosene* category was more difficult. Training was not 100% correct and the size of the network required was too large. The ratio of samples to varaibles in the network for this step was lower than desired, ie. less than one. Information about the compounds used to make the clasifications and architecture of the networks trained is collected in Tables 5 and 6.

Table 5: Compounds Used in Classification

Compound	Step 1: Broad Categories	Step 2: *Special* Fuels	Step: 2: *Kerosene* Fuels
1,2,3,4-Tetrahydronaphthalene	x	x	x
Naphthalene	x		x
2-Methylnaphthalene	x		
1-Ethylnaphthalene	x		x
1,3-Dimethylnaphathalene		x	
2,3- & 1,4-Dimethylnaphathalene	x		x
1,8-Dimethylnaphathalene			x
1,4,6,7-Tetramethylnaphthalene	x	x	x

Table 6: Network Architecture

Architecture Property	Step 1: Broad Categories	Step 2: *Special* Fuels	Step: 2: *Kerosene* Fuels
Input nodes	6	3	6
Hidden nodes	5	2	5
Output nodes	3	3	3
Variables	53	17	53
Samples	104	39	52
Ratio Samples:Variables	1.96	2.29	0.98
% Correct Classification	100%	100%	98%

The approach involving first combining JetA and JP-8 fuels into a single class and then classifying those fuels in a second step was more successful. Networks trained using the concentrations of naphthalene and naphthalene derivatives as input patterns were again too large or could not be trained to 100% correctly classify fuels. However, networks trained using the concentrations of six benzene- and naphthalene-related compounds were trained to 100% correctly classify the fuels. The optimum architecture included six input, six hidden, and six output nodes for the first step, and six input, three hidden, and two output nodes for the second step. These networks were small enough that the ratio of samples to variables was appropriate, ie. greater than one. The list of compounds used to make the classifications is collected in Table 7.

Table 7: Compounds Used in Classification

Isopropylbenzene

1,2-Diethylbenzene

1,2,3,4-Tetramethylbenzene

1,2,3,4-Tetrahydronaphthalene

1-Ethylnaphthalene

2,3-Dimethlynaphthalene

The results of the validation process are collected in Table 8. It is important to keep in mind the four training/prediction set (TSET/PSET) pairs are composed of *fresh* fuels. While the results were good for the classification of *fresh* fuels as shown in Table 8, the results for a prediction subset of *weathered* fuels were not good. This was disappointing given one goal of this research was to be able to correctly

samples of *weathered* fuels. However, the reason becomes clear when the composition of the *weathered* fuels is examined.

Figure 2. indicates how the concentrations of the compounds used classification change in a sample of JP-8 as it evaporates. Similar changes are observed in samples of other fuels. It seems unlikely that the classification scheme will be successful if the composition of these components change in this fashion. Clearly, further efforts should be made to quantify additional compounds that are less likely to be volatile.

Table 8: Training and Prediction Results

Step 1

Class	All Four Training Subsets		Prediction Subset 1		Prediction Subset 2		Prediction Subset 3		Prediction Subset 4	
	Number	Score	Number	Score	Number	Score	Number	Score	Number	Score
JP-4	18	100%	2	100%	2	100%	2	50%	2	100%
JP-7	7	100%	1	100%	1	100%	1	100%	1	100%
JPTS	9	100%	1	100%	1	100%	1	100%	1	100%
JP-5	11	100%	1	100%	1	100%	1	100%	1	100%
JP-8 & JetA	36	100%	4	75%	4	100%	4	100%	4	100%
AvGas	11	100%	1	100%	1	100%	1	100%	1	100%
Total/Average	92	100%	10	90%	10	100%	10	90%	10	100%

Step 2

Class	Training Subsets 1, 2 &		Training Subset 3		Prediction Subsets 1 & 3		Prediction Subset 2		Prediction Subset 4	
	Number	Score	Number	Score	Number	Score	Number	Score	Number	Score
JetA	18	100%	19	100%	2	100%	2	50%	2	100%
JP-8 & JetA	18	100%	17	100%	2	100%	2	50%	2	50%
Total/Avera	36	100%	36	100%	4	100%	4	50%	4	75%

Figure 2.

Concentrations in a JP-8 as it Weathers

11 - 11
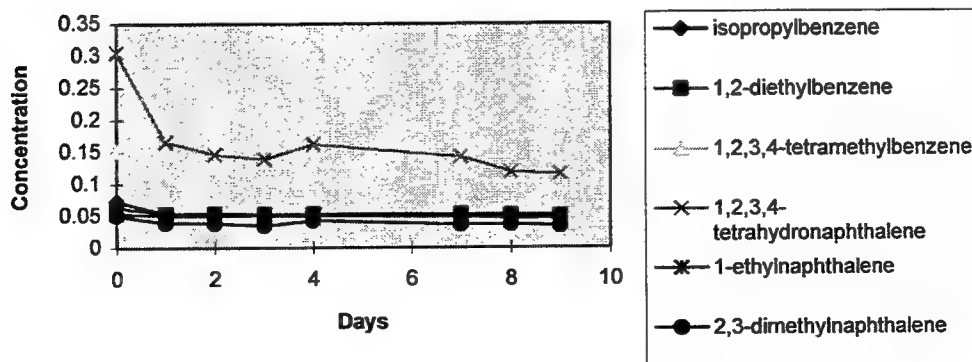
Part II.

Introduction

A second approach to the problem of identifying fresh and *weathered* aviation fuels involves the application of vapor sensitive detectors and artificial neural network analysis. Arrays of vapor sensitive detectors can be constructed to respond in a measurably different manner to samples of pure compounds and mixtures [12]. Such arrays are of preactical interest because they can be microprocessor controlled and small enough to be easily portable [13]. They have been used to detect and identify a number of hazardous vapors [13,14]. A review of chemical sensors by Janata, Josowicz and DeVaney [15] includes references on the use of sensors for detection of odors, gases of environmental interest and the development of an *electronic nose.*

Previous work by the author in collaboration with this laboratory reported on the identification of aviation fuels by neural network analysis of the response of an array of commercial sensors containing metal oxides [18]. The metal oxide in each sensor is embedded with electronic heaters and collectors, and its resistance is a function of the gases or vapors in its environment. Any change in the resistance is detected, processed and used in the identification of the vapor.

An array of n detectors will generate n signals, in effect a data vector of length n. The preliminary study reported here probed the similarity of the response of an array of vapor sensitive detectors to fresh and *weathered* fuels simply by comparing the the data vectors for fresh and *weathered* samples of different classes of aviation fuels.

Experimental Section

Overview of the Experiment: The general set-up of experiment as displayed in Figure 3 was essentially the same as previously reported [18]. Samples of fuel (2 μL) were injected into a stream (15 mL/min.) of carrier gas through the injection port from a gas chromatograph. The injection port was used simply to get the liquid sample into the vapor phase and into the stream of carrier gas. No chromatography was done in these experiments. The aspirated sample was swept into the sample chamber through Teflon tubing and

exposed to an array of eight (8) vapor sensitive metal oxide sensors. The analog response of the sensors was digitized at one second intervals one thousand times and stored for processing.

Figure 3. General set-up of the experiment
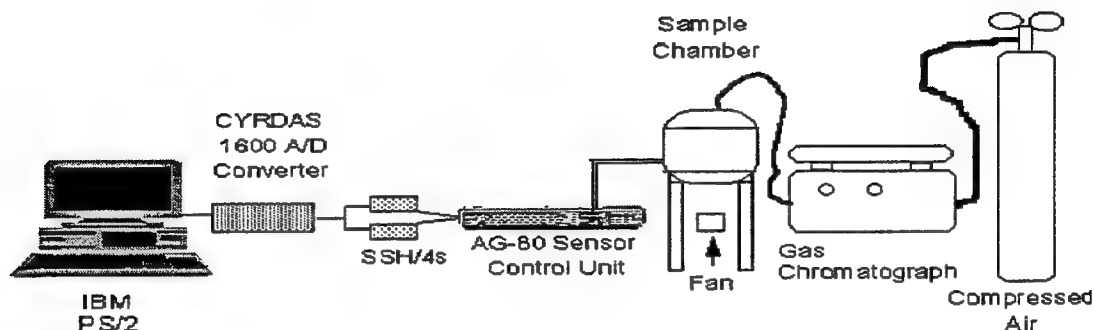
Data Set: Samples of the *weathered* fuel were drawn from the same pool and prepared described in Part I of this report. Two samples taken from a monitoring well number at a fire training area that has been cloased since 1992. One sample was drawn from the well in September, 1995; the other was drawn in July, 1996. These samples are stored in a refrigerator when not being used.

Chamber: The sensors were mounted in a sample chamber fabricated from schedule 40 PVC pipe and PVC endcaps. The approximate volume of the chamber was 20 Liters. Sample entered the chamber through a Teflon tube mounted on the inside of the top endcap. The inlet tube pointed directly toward the center of a mixing propeller mounted at the bottom of the chamber. The mixing propeller was connected to a fan motor by a shaft passing through the bottom endcap. The arrangement of the sensors in the chamber is displayed in Figure 4. Four small vent holes in the top of the chamber allowed sample and carrier gas to escape. One more tube was mounted through the top of the chamber so an additional stream of air could be used to sweep sample out of the chamber after each run.

Figure 4. Orientation of the sensors in the sample chamber.
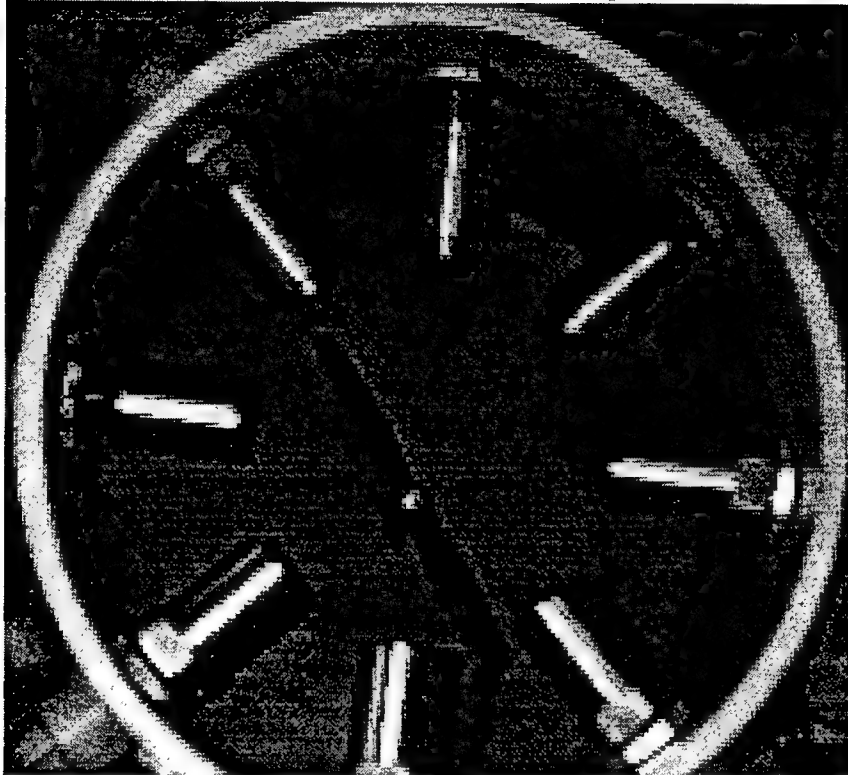
Sensors: The sensors and sensor controllers were purchased from International Sensor Technology, Irvine, CA. The eight sensors purchased were described as being for benzene, diesel fuel, ethyl benzene, fuel oil, gasoline, JP-4, JP-5, and toluene. The ethyl benzene sensor was set at the factory to respond to concentrations in the range 0 to 200 ppm; the fuel oil sensor was set to respond to concentrations of 0 to 100 ppm. The other six sensors were set to respond to concentrations in the range 0 to 1000 ppm. The sensors were controlled by an AG80 multi-channel control unit. The response of each channel was available from the control unit as an analog output signal in the range 0 to 1 volt DC. The control unit also had an analog meter that could be used to monitor visually the progress of a run and the sweeping out of the chamber afterwards.

Data Collection: The analog signal from each of the eight sensors was digitized one thousand times at one second intervals, digitized and stored. An IBM PS/2 model 30 personal computer sampled the analog output of the AG80 sensor control unit through a Model No. 1601 CYRDAS 1601 analog-to-digital converter card and two Model SSH/4 sample-and-hold boxes from CyberResearch, Inc., Branford, CT.

11 - 14

Each channel of output from the AG80 was connected to a channel on a sample-and-hold box, and the boxes daisy-chained so that all eight channels of output could be digitized almost simultaneously. The data from each run was stored as a text file in the form of an 1000x8 matrix containing 1000 response for each of the eight sensors. The software used to accomplish the data acquisition was the DAS-1600 Test and Control Panel provided with the A/D card running under DOS Version 3.30.

Data Processing: The 1000x8 data matrix from each run was analyzed on an ABC Products 80386sx Notebook Sized Computer using Quattro Pro 4 [16] and PS-PLOT Version 1.12 [17] running under Windows 3.1. The response of the sensors *grew in* as sample was swept from the injection port to and diffused throughout the sample chamber. The response of each sensor to a sample was taken as the average of the measurements over the range it was stable. For this study, that range was taken to be from measurement 650 to 1000. The eight average responses were stored as a 1x8 data vector representing the responses of each of the eight sensors to a single fuel. This 1x8 data vector represented the *pattern* of the response of the array to each sample. Duplicate runs were made for each sample to check reproducibility.

Results and Discussion

The final results of this preliminary study are collected in Table 9 The 1x8 data vector was reduced to a 1x5 data vector based on previous experience with these detectors. The response of the benzene sensor was removed because it is almost always zero. The responses of the diesel, gasoline and JP-4 sensors generally exhibit high pairwise correlations. Therefore only one of those responses, that of the diesel sensor, is shown in the table.

One sample picked at random for the *weathering* study turned out to be the sample described in Part 1 which we believe to be mislabeled. The response of the array to this sample was different from the response of the array to an authentic sample of JP-8.

It must be said that the data is not as reproducible as one might wish. Clearly, an arrray must be found that is more stable. However, it is surprising that the response of the arrray to the *weathered* fuel is as similar to that of the *fresh* fuel as it is. The results for the two samples drawn from the monitoring well

at the former fire training area are particularly intriguing. The response of the array to these samples is

very similar to that of the array to the sample we believe to be a JP-4 mislabeled as a JP-8.

Table 9: Sensor Responses to Fresh and Weathered Fuels

Sample	Diesel	Ethyl benzene	Fuel Oil	JP-5	Toluene	Comments
F2317.1	0.000	0.076	0.066	0.000	1.013	A mislabeled JP4
F2317.2	0.000	0.044	0.077	0.000	0.989	Shakey ... but
F2317.3	0.000	0.107	0.085	0.000	1.017	Best! - fresh "phony" JP-8
W2317.1	0.000	0.081	0.057	0.000	1.014	Weathered Sample
W2317.2	0.000	0.071	0.078	0.000	1.008	weathered "phony" JP-8
W2317he.3	0.000	0.000	0.000	0.000	0.415	weathered "phony" JP-8
W2317he.4	0.000	0.101	0.000	0.000	0.985	weathered "phony" JP-8
F2030.1	0.000	0.259	0.098	0.000	1.011	fresh JP-8
F2030he.1	0.000	0.316	0.092	0.000	1.001	fresh JP-8 - modified set-up
W2030he.1	0.000	0.320	0.017	0.000	0.816	weathered JP-8
W2030he.2	0.000	0.205	0.000	0.000	0.710	
W2030he.3	0.000	0.318	0.047	0.000	0.891	
F152he.1	0.000	0.135	0.000	0.000	0.603	fresh Jet-A
W152he.1	0.000	0.167	0.000	0.000	0.405	weathered Jet-A
W152he.2	0.000	0.177	0.000	0.000	0.573	
F012he.1	0.000	0.205	0.000	0.000	0.742	fresh JP-5
W012he.1	0.000	0.248	0.000	0.002	0.633	weathered JP-5
W012he.2	0.000	0.079	0.000	0.000	0.326	problems - flush air on at star
W012he.3	0.000	0.261	0.000	0.000	0.744	
F2013he.1	0.000	0.000	0.000	0.000	0.747	fresh JPTS - septum?
F2013he.2	0.000	0.000	0.000	0.000	0.776	new septum
F2013he.3	0.000	0.000	0.000	0.000	0.789_	
F2013he.4	0.000	0.006	0.001	0.000	0.945	
W2013he.1	0.000	0.054	0.042	0.000	0.886	weathered JPTS
W2013he.2	0.000	0.023	0.000	0.000	0.963	
W2013he.3	0.000	0.056	0.016	0.000	0.995	
W2013he.4	0.000	0.000	0.000	0.000	0.839	what the heck?
W2030he.4	0.002	0.279	0.004	0.009	0.844	
W2013he.5	0.000	0.109	0.104	0.000	1.020	
MW5-95.1	0.000	0.086	0.000	0.000	1.007	MW5 - 1995
MW5-95.2	0.000	0.090	0.000	0.000	1.006	MW5 - 1995
MW5-96.1	0.000	0.101	0.000	0.000	1.009	MW5 - 1996
MW5-96.2	0.000	0.194	0.012	0.000	1.009	MW5 - 1996

Disclaimer

Certain commercial equipment, instruments and materials are identified in this report to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Acknowledgments

<u>Bibliography</u>

1. Zemo, D.A.; Bruya, J.E.; Graf, T.E. *Ground Water Monitoring & Remediation,* **1995**, *Spring,* 147-156.

2. Long, J.R.; Mayfield, H.T.; Henley, M.V.; Kromann, P.R. *Anal. Chem.,* **1991**, *63,* 1261-1264.

3. Faruque, A., Lavine, B. K., and Mayfield, H.T., <u>FIP: A Pattern Recognition Program for Fuel Spill Identification,</u> USAF Report AL/EQ-TR-1996-0007, Tyndall AFB, FL, May 1996, Unclassified (Available through DTIC).

4. Lavine, B.K.; Qin, X.; Stine, A.; Mayfield, H.T. *Process Control Quality,* **1992**, *2,* 347-355.

5. Lavine, B.K.; Stine, A.; Mayfield, H.T. *Anal. Chimica Acta.* **1993**, *277,* 357-367.

6. Neural Network TOOLBOX: For Use with MATLAB^{tm}, The Mathworks, Inc. Natick, Massachusetts, 1992.

7. Andrews, J.M.; Liberman, S.H. *Anal, Chim. Acta.* **1994**, *285,* 237-246.

8. McCarrick, C.W.; Ohmer, D.T.; Gilliland, L.A.; Edwards, P.A.; Mayfield, H.T. in review for publication in *Anal. Chem.,* **1996**.

9. Wythoff, B.J. *Chemom. Intell. Lab Syst.* **1993**, *18,* 115-155.

10. MATLAB^{tm}, High-Performance Numerical Computation and Visualization Software, The Mathworks, Inc. Natick, Massachusetts, 1992.

11. In: *Handbook of Aviation Fuel Properties.* Coordinating Research Council, Inc., Atlanta, GA, 1983, p. 2-18. (Available from: Society of Automotive Engineers, Inc., General Publications Department, 400 Commonwealth Drive, Warrendale, PA, 15096).

12. Stetter J.R. Electrochemical Sensors, Sensor Arrays, and Computer Algorithms for Detection and Identification of Airborne Chemicals, In: *Fundamentals and Application of Chemical Sensors:* Schuetzle, D; Hammerle, R.; Ed. ACS Symposium Series, 300, American Chemical Society, Washington, DC. 1986, p. 299-308.

13. Stetter, J.R.; Jurs, P.C.; Rose, S.L. *Anal. Chem.* **1986**, *58,* 860-880.

14. Rose-Pehrsson, S.L.; Grate, J.W.; Ballantine, D.S.;Jr; Jurs, P.C. *Anal. Chem.* **1988**, *60,* 2801-2811.

15. Janata, J.; Rosowicz, M.; Devaney, D.M. *Anal. Chem.* **1994**, *66,* 207R-228R.

16. Quattro Pro 4, Borland International, Inc., 1992.

17. PS-Plot Version 1.12, Polysoft Ltd., 1992.

18. McCarrick, C. W., Ohmer, D. T., Gilliland, L. A., Edwards, P. A., and Mayfield, H. T., *Anal. Chem.*, **1996**, in press.

19. Ballantine, D.S., Jr.; Rose, S.L.; Grate, J.W.; Wohltjen, H. *Anal. Chem.*, **1986**, *58*, 3058-3066.

20. Zellers, E.T.; Batterman, S.A;, Han, M.; Patrash, S.J. *Anal. Chem.*, **1995**, 67, 1092-1106.

21. Brown, S.D.; Blank, T.B.; Sum, S.T.; Weyer, L.G. *Anal. Chem.* **1994**, *66*, 315R-359R.

22. Brown, S.D.; Sum, S.T.; Despagne, F.; Lavine, B.K. *Anal. Chem.* **1996**, *68*, 215R-61R.

23. Duin, R.P.W. *Pattern Recognit. Lett.* **1994**, *15*, 215-217.

24. Windows for Workgroups Version 3.1, Microsoft Corporation, 1994.

25. Excel Version 5.0C. Microsoft Corporation, 1985-1994.

A STUDY OF OXIDATIVE REACTIONS MEDIATED
BY LASER-EXCITED OCULAR MELANIN

Randolph D. Glickman
Associate Professor
Department of Ophthalmology

University of Texas Health Science Center at San Antonio
7703 Floyd Curl Drive
San Antonio, TX 78284-6230

A STUDY OF OXIDATIVE REACTIONS MEDIATED
BY LASER-EXCITED OCULAR MELANIN

Randolph D. Glickman, Ph.D.
Associate Professor
Department of Ophthalmology
University of Texas Health Science Center at San Antonio

Abstract

Melanin is excited to a free radical when exposed to visible light. Normally, the reactivity of the melanin radical is contained by the melanosome complex. After laser injury to ocular tissue, however, the possibility exists that melanin in damaged melanosomes of the retinal pigment epithelium (RPE) could react with physiological substrates to produce photochemical damage in surrounding tissue. The ability of the melanin radical to oxidize nicotinamide adenine dinucleotide phosphate (NADPH) during light exposure was used to characterize the reactivity of RPE melanosomes after various experimental treatments. Melanosomes disrupted by the fundamental output of the Nd:YAG laser oxidized significantly more NADPH than did intact ones during subsequent exposure to a visible light source. As the sample irradiance produced by the YAG laser exposure went up, melanosome photoreactivity increased. The rate of increase was used to estimate that the threshold for melanosome disruption by a 10 nsec pulse at 1064 nm was about 560 mJ/cm^2. Based on the amount of NADPH oxidized by melanosomes excited by various laser wavelengths, the action spectrum of the melanin radical has a broad peak in the visible spectrum between 450 and 500 nm. The melanin radical was also detected with two fluorescent probes sensitive to oxidation, 2',7'-dichlorofluorescin (DCFH) and dihydrorhodamine 123 (DHR123). Both probes showed a progressive increase in fluorescence when mixed with RPE melanosomes and excited with the 496 nm line from an Argon ion laser. The growth in the fluorescence signal increased with the number of melanosomes in the reaction mixture, laser excitation power, and prior photodisruption of the melanosomes. Addition of ascorbic acid, an aantioxidant, quenched the fluorescence. DHR123 was less sensitive to AA quenching than was DCFH, possibly indicating a lower oxidation potential than that of AA. The ability of these fluorescent probes to detect melanin radical activation may be useful in investigating mechanisms of tissue damage in the retina and RPE following short pulse laser exposures.

Introduction

Although the biological pigment melanin is found in diverse tissues such as hair and CNS, its most active role may be in epithelial tissues of the skin and eye where it is generally thought to play a protective role against the effects of optical radiation. For example, in the eye, up to 60% of all light entering the pupil is absorbed by the melanin contained in the retinal pigment epithelium (RPE) and choroidal tissues[1]. The active mechanisms of melanin have not been appreciated until recently. Historically the major role of melanin has been thought to be bulk absorption of light which sharpened vision by reducing intraocular light scatter; other biological roles have been assumed to be unimportant[2]. In contrast, melanin has been proposed as a free raducal scavenger, thus helping to prevent photochemical damage[3,4].

More recently, the possibility that melanin may be able to support or promote free radical reactions has been recognized. Although it has long been known that natural and synthetic melanin is excited to a free radical by optical radiation[5,6], the significance of this observation was not known. Indeed, the melanin radical has been described as "stable"[7] and thought to be incapable of significant reactions or, at least, isolated in its intracellular locations thus preventing interactions with cellular components. Arguing against the nonreactivity of melanin are the many observations of Sarna and his colleagues that melanin under illumination of UV or visible light is able to evolve reactive oxygen species[8,9]. Although these experiments were largely carried out under non-physiological conditions and the relevance to the intracellular milieu was uncertain, the results at least raised the possibility that melanin was as likely to promote light damage in the cell as to prevent it. The reaction of photoactivated melanin with physiological substrates such as NADPH[10,11] and ascorbic acid[12,13] indicated that, under appropriate conditions, melanin radicals induced by the absorption of light in ocular tissue may promote photochemical damage.

Melanin, by virtue of being a broadband absorber, is involved in thermal and photodisruptive damage to the RPE and surrounding tissues produced by continuous wave (CW) and pulsed lasers. Indeed, it is the involvement of melanin in thermal and photodisruptive processes initiated by laser irradiation of the eye which inspired the current investigation. The heating processes in the melanin granules have been described in several models of thermal damage to the eye, most recently by Thompson et al.[14]. In the case of shorter laser pulses, in the nanosecond regime and below, the role of melanin may be complex. The initial effect of a short pulse laser exposure will be to disrupt the melanin granule (melanosome) through local thermal transients leading to explosive vaporization, or because of photomechanical processes.

The precise mode of tissue damage by ultrashort pulses is still under investigation[15,16].

Determinations of tissue damage thresholds by picosecond and shorter laser pulses[17] suggested that at 90 fs, there was a delayed appearance of a lesion, for which no explanation was immediately apparent. Taking into consideration the slowly developing nature of the damage, and also the known increase in photoreactivity of disrupted melanosomes[18], Glickman[19] suggested that photochemical damage promoted by laser-disrupted melanosomes might be responsible for the slowly developing lesions. In order to test this hypothesis, some means of detecting photo-oxidative reactions in situ or *in vivo* are required. This summer project was undertaken to characterize more fully the photochemical reactions of intact and disrupted melanin, including the action spectra of such reactions, and to determine if melanin-mediated oxidative reactions could be detected with fluorescent probes in an *in vitro* model system.

Methods

Preparation of Melanosomes. Melanosomes were isolated from the RPE cells of freshly obtained bovine eyes as previously described[12,13]. Melanosomes were stored frozen at $-20°$ C in 0.25 M sucrose until use.

Assay for NADPH Photooxidation. The oxidation of β-nicotinamide adenine dinucleotide phosphate (NADPH) was used as a biochemical marker of melanin activation[18]. Assays were conducted in 1.5 ml microcentrifuge tubes and contained 7 mM NADPH (Sigma Chemical Co., #N1630), 80 mM Tris buffer (pH 7.4), and various dilutions of bovine RPE melanosomes (ranging from 6,240 to 312,000 granules/assay), in a total reaction volume of 25 μl. Generally, sufficient reagents were mixed to make up a pair of samples; one member of the pair was exposed to the experimental treatment, and the other served as the control. The determination of NADPH oxidation was accomplished spectrophotometrically. The reduced form of NADPH has an absorption peak at 340 nm, while the oxidized form does not. Prior to spectrophotometry, the reaction mixtures were diluted with 1.5 ml of deionized H_2O and passed through 0.45 μm pore syringe filters (Sun Brokers, Wilmington, NC) to remove the melanosomes. With the samples in quartz cells, the absorbance at 260 and 340 nm was measured on a Beckman DU-64 spectrophotometer. Using an extinction coefficient for NADPH of 6401 cm^{-1}M^{-1} (empirically determined from a fresh solution of NADPH), the loss of absorbance at 340 nm compared to control samples was converted into μmoles of NADPH oxidized.

Photochemical Activation of Melanosomes by Continuous Wave Laser Exposure. Free radical activity was induced in reaction mixtures of melanosomes and NADPH by a 5 min exposure to the 488 nm output of a Argon ion CW laser (Coherent Laser Innova 200), delivering 150 mW into a 3 mm beam. In earlier experiments, the laser energy was conducted into the sample with

a fiber optic, while in later experiments the laser beam was steered directly into the sample with mirrors, wih equivalent results. Sample irradiance was measured with a Molectron EPM-500 with a PM-10 head. Following CW laser exposure, the samples were processed for NADPH oxidation.

Photodisruption of Melanosomes by Q-Switched Nd:YAG Laser Exposure. A Spectra-Physics DCR-11 operating in fundamental mode (1064 nm at a nominal pulsewidth of 10 nsec full width half maximum) was used to produce the pulsed laser energy that disrupted the melanosomes. The laser was fitted with Gaussian optics to produce a beam with a relatively uniform energy distribution. Beam size was estimated from the burn pattern produced on laser "burn" paper (Kentek Co.). Pulsewidth was measured with a Scientech #350 photodiode with the output displayed on a Tektronix DSA-602 digitizing signal analyzer. The beam was then expanded (to reduce the probability of damaging the steering and control optics) and passed through a polarizer and fresnel rhomb, which allowed adjusting the energy per pulse reaching the sample by the registration of the two optics. The beam was re-focussed to a nominal 1.9 mm diameter beam into the sample. Laser radiometry was performed with a Molectron JG-2000 ratiometer with J25 joule (pulse) detectors. One of the detectors (at the reference position) measured a portion of the beam deflected by a beamsplitter, and the other was placed at the sample position. Once the ratiometer was calibrated, the detector at the sample position was removed and the energy per pulse calculated from the reference detector. To induce photodisruption, the melanosomes in the standard NADPH reaction mixture were exposed to 1800 pulses from the Nd:YAG laser at 10 pps. Following the YAG laser exposure, the sample was "pumped" with the Ar laser for the standard 5 min period, and then NADPH oxidation was measured.

Determination of Action Spectrum for the Photooxidation of NADPH by Melanin. Thirteen wavelengths between 406 .7 nm and 647.1 nm were used to pump standard reaction mixtures of NADPH, melanosomes, and Tris buffer. These lines were all, with one exception, tuned from an Ar (Coherent Innova 200) or a Kr (Coherent Innova 100) CW laser. The one exception was the 441.6 nm line which was obtained from two HeCd (Omnichrome Series 56) lasers with the beams ganged. This was done to make the sample irradiance more comparable to that achieved with the Coherent lasers; however, it was still less, and thus the data obtained at 441.6 nm is less reliable than at the other wavelengths tested. The sample irradiance was adjusted at each wavelength to deliver the same photon flux into the sample, i.e. ~3.18 x 10^{21} photons/cm^2 for 300 sec. Because this calculation depended critically on the laser beam size at the sample, the beam diameter was measured by a classical knife edge scan (the beam width between 14% and

86% of maximum power), and with a video-based beam profiler (Coherent Laser BeamView PC). The discrepancy between the two measurement was less than 5% (often less than 1%). Following the laser exposure, the amount of NADPH oxidized was determined.

Preparation of Fluorescent Probes. 2',7'-dichlorofluorescin diacetate (oxidized form is -scein, F.W. = 487.29), and dihydrorhodamine 123 (F.W. = 346.38) were purchased from Molecular Probes of Eugene, OR (product numbers D399 and D632, respectively). The preparation and use of these probes were adapted from the procedures of LeBel et al.[21] and Royall and Ischiropoulos[22]. Ten mg of 2',7'-dichlorofluorescin diacetate was dissolved in 20 ml of methanol to yield nominal 1 mM stock solution. This was kept stored at -20° C under N_2 gas. The diacetate ester was hydrolyzed to free 2',7'-dichlorofluorescin (DCFH) before use by mixing 0.5 ml of the diacetate stock with 2.0 ml of 0.01 N NaOH and incubating for 30 min at room temperature. The hydrolysis was stopped by adding 10 ml of 25 mM NaH_2PO_4 (pH 7.5) to the reaction. This working solution was usually prepared fresh for each day's work, and stored on ice in the dark prior to use. Dihydrorhodamine 123 (DHR123) was supplied as 10 mg of a dry powder in a glass ampoule. This was dissolved in 1 ml of N,N-dimethylformamide. This produced a stock solution of 28.9 mM, which was transferred to a small, glass bottle, sparged with N_2, and stored at -20° C in a desiccator.

To make the working dilution of DCFH, typically 50 µl was combined with variable amounts of melanosomes, Tris buffer, and other reagents such as ascorbic acid, to a total volume of 1.5 ml. This brought the working concentration of DCFH to 1.4 µM, and the residual methanol to 0.133% (v/v). In the case of DHR123, 5µl of the stock solution of the probe was combined with Tris or water and other reagents to a total volume of 1.5 ml. The working concentration of DHR123 was 5 µM, and the residual dimethylformamide was 0.017% (v/v). Dimethylformamide, even at this low concentration, precipitated biological buffers used in the reaction mixtures, and also produced a cloudy appearance with water alone. Comparative experiments with Tris, HEPES, and water alone indicated that, in the *in vitro* system, a buffer was not necessary, and comparable fluorescence readings were obtained with water alone as the diluent.

Fluorescence Measurements. The reaction mixtures were contained in quartz fluorimeter cells with all-polished walls (Starna cell #3-Q-10). The cells were capped and stirred with a micro stir-bar apparatus (Starna model 9400) during measurements. Fluorescence readings and spectroscopy were made with a SPEX 1404 double spectrometer with a cooled photomultiplier photon detector. The slit width of both monochromators was set to 100 µm. A 2 in diameter, 38 mm focal length lens was placed 5.4 cm from the sample cell and 24.5 cm from the entrance

slit of the SPEX. The final configuration of cell and lens was set empirically by adjusting the sample position to maximize the counts per second (CPS) measured from a fluorescing sample. Excitation of the samples was usually with the 496 nm line from the Coherent Innova 200 Ar laser in a 3 mm beam into the cell. Fluorescence data was recorded on computer, later being converted into ASCII files for analysis and graphical output.

Results

Photodisruption Threshold for RPE Melanosomes, 1064 nm. RPE melanosomes will oxidize the dinucleotide cofactor, NADPH, when excited by optical radiation. The extent of NADPH oxidation depends on the wavelength and intensity of the radiation (vide infra), thus the amount of NADPH oxidized serves as a marker for the number of free melanin free radicals induced in the melanosomes. Previous work[18] established that disrupting the melanosomes with Q-switched laser exposures increased their photoreactivity, presumably by making melanin radical sites in the interior of the melanosome accessible to the substrate. Scanning electron microscopy of the laser-disrupted melanosomes supported this hypothesis by revealing major fissures and splits in the melanosomes[18]. The exposures in the previous work were made with the frequency-doubled output of the Nd:YAG laser (532 nm @ 10 ns nominal pulsewidth). Based on the NADPH oxidation obtained, a threshold for melanosome disruption was extrapolated and reported to be about 150 mJ/cm². In the present experiments, we duplicated the experimental procedure, but used the fundamental Nd:YAG output, i.e. 1064 nm @ 10 ns nominal pulsewidth for melanosome disruption.

Because the absorption coefficient of melanin in the NIR is thought to be 5-10 times lower than that in the visible[20], the threshold irradiance for melanosomes disruption should be correspondingly higher for the fundamental YAG exposures.

After the melanosomes were disrupted, the magnitude of difference in NADPH oxidation depended on the duration of the "pump" time by the CW Ar laser. This period of irradiation excited the melanin radical and caused the NADPH to be oxidized. Without "pumping", there should be little difference between

NADPH Photooxidation
YAG Exposed vs Non-Exposed Melanosomes
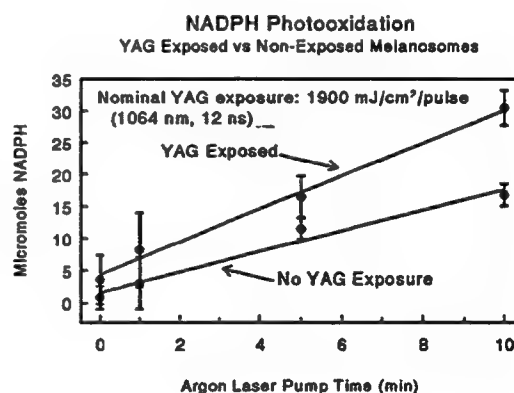
Fig. 1. NADPH oxidation by native (intact) and photodisrupted bovine melanosomes. YAG exposure was 1800 pulses at 10 pps. Argon pump time refers to the duration of the CW exposure which followed the YAG exposure.

12-7

disrupted and native melanosomes in the amount of NADPH oxidized. The first experiment demonstrated that this was in fact so, and that a difference between the two only emerged after the Ar laser exposure (Fig 1).

Standardizing the Ar laser pump time to 5 minutes at 2.12 W/cm² allowed comparison of NADPH oxidation between melanosome samples exposed to increasing energy per pulse from the YAG laser. This data set is shown in Fig. 2. Although the data are rather noisy, extrapolation to zero NADPH oxidation pro- duced a threshold estimate between 500 and 600 mJ/cm², while the actual zero data values occurred at a sample irradiance of 800 mJ/cm²,

Photodisruption Threshold
YAG laser, 1064 nm, 10 ns pulses

Fig. 2. Estimation of melanosome disruption threshold by the amount of NADPH oxidized in a standard assay. Melanosomes disrupted by 1800 pulses at 1064 nm from a Nd:YAG laser. Disruption threshold estimated at between 500 and 800 mJ/cm².

which is reasonably close to the anticipated range (750 to 1500 mJ/cm²).

Action Spectrum of Melanin Radical Estimated From NADPH Oxidation. Using the amount of NADPH oxidized as an indirect measure of melanin radical excitation, an action spectrum was constructed over the range of 406.7 to 647.1 nm, using the visible wavelengths available from

the Ar, Kr, and HeCd CW lasers (Fig. 3). All the exposures with the exception of the 441.6 nm from the HeCd laser were adjusted to deliver 3.18×10^{21} photons/cm² to the sample. The power available from the HeCd laser was insufficient to reach this level, so the data obtained with this source were scaled up, assuming a linear relationship between sample irradiance and NADPH oxidation. This maneuver increased the variance in the data, as reflected in the large error bars for data point at 441.6 nm in Fig. 3. However, the rest of the data is fairly consistent and reveals that the efficiency for exciting the melanin radical is greatest for wavelengths between 450 and 500

Photooxidation of NADPH by Melanin
Visible Action Spectrum

Fig. 3. Action spectrum of NADPH oxidation by bovine melanosomes excited by visible wavelengths. All exposures were equated to deliver 3.18×10^{21} photons/cm².

nm, and decreases at shorter and longer wavelengths.

As a demonstration of the physiological significance of the melanin radical action spectrum, an experiment was performed to characterize the photooxidation of retinal proteins mediated by melanin. A crude retinal protein extract was prepared from a sample of baboon retina obtained from the Pathology Laboratory of the Southwest Foundation for Biomedical Research, San Antonio, TX, and mixed with an aliquot of bovine melanosomes. These mixtures were then exposed to the 488 nm line of the Ar ion laser at 1.052 W/cm² or to the 568 nm line of the Kr ion laser at 0.894 W/cm² for 10 min. These irradiances were set to deliver an equivalent number of photons to the samples. At the end of the exposure time, the proteins were separated by polyacrylamide gel electrophoresis and then transferred in a Western blot to nitrocellulose membrane. These blots were processed with the Oxyblot kit (Oncor Inc., Gaithersberg, MD), which produced a chemiluminescent (CL) signal for oxidized proteins. The kit actually detected carbonyl adducts which resulted from hydroxyl radicals attacking proteins. The light produced by the CL reactions exposed a sheet of X-ray film placed over the Western blots, producing bands of exposed film corresponding to the locations of the protein bands in the original gel. The films were scanned in a Beckman DU-640 spectropho-

Photooxidation of Retinal Protein
Comparison of 488 and 568 nm laser

Fig. 4. Protein oxidation by laser-excited melanin granules. "+488" is the Ar+ ion, and "+568" is the Kr+ ion, laser outputs. Carbonyl adducts on proteins detected by chemiluminescence assay on Western blots of baboon retinal proteins. Molecular weight standards indicated at top of graph. See text for further details.

tometer equipped with a densitometry accessory, and the film-scan data are shown in Fig 4. It is clear that, for photon-equivalent exposures, the blue (488 nm) light from the Ar laser is more efficient than the yellow (568 nm) light from the Kr laser in inducing melanin-mediated oxidation, at least in the 48-66 KDa proteins. This is consistent with the action spectrum determined with NADPH oxidation. It was also found that the antioxidant, glutathione, was effective in preventing protein oxidation in this experiment (data not shown).

Detection of Melanin Photooxidative Reactions with Fluorescent Probes. The focal point of the work this summer was to determine if melanin radical activation could be detected optically with a fluorescent probe. Two commercially available probes were tested: 2',7'-dichlorofluorescin (DCFH) and dihydrorhodamine 123 (DHR123). These probes have the

property that when they are chemically reduced they are non-fluorescent, but when oxidized are highly fluorescent. Both probes have excitation and emission maxima in the visible spectrum (vide infra). Because of time limitations in this summer project, extensive methods development could not be done. Therefore, we chose to follow published methods for the preparation and use of these particular probes[21,22]. These methods proved to be satisfactory and light-activated melanin oxidized both DCFH and DHR123.

The first experiments determined the optimal wavelength for fluorescence excitation. A sample of DCFH was oxidized by a chemical system (ferrous sulfate and hydrogen peroxide[21], that generated OH· radicals principally by the following reaction:

$$Fe^{+2} + H_2O_2 \longrightarrow Fe^{+3} + OH· + OH^- \tag{1}$$

The oxidized DCFH was then excited by various wavelengths from the Ar ion laser (488, 496.5, 501.7, and 514.5 nm) and the maximum amplitude of the fluorescence was determined. The results are shown in Fig 5. The peak excitation wavelength was around 500 nm, in accordance with the value of 502 nm reported by Royall and Ischiropoulos[22]. It was noted that the spectrum of the fluorescence did not markedly change with the different excitation wavelengths. The 496.5 nm line was selected as the excitation wavelength for the rest of the experiments because the Ar ion laser emitted more power at that line than at the 501.7 line.
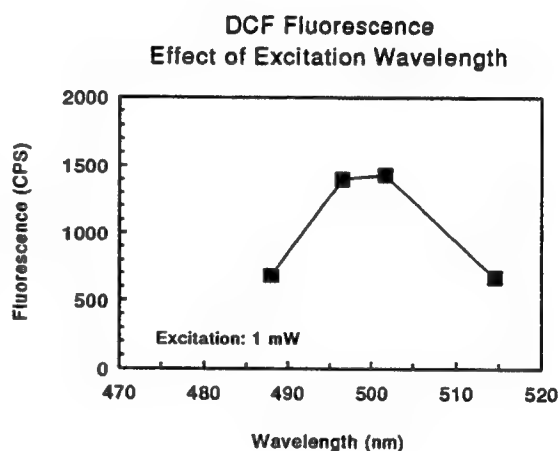
Fig. 5. Determination of peak wavelength (abscissa) for excitation of DCFH fluorescence (ordinate).

This wavelength was also near the 500 nm optimum reported for DHR123[22] .

Autooxidation in room air of DCFH was measured by placing the probe at its working concentration in a cell with the cap on and stir-bar engaged, and measuring the fluorescence at 10 min intervals with the excitation laser beam set to 1 mW. The laser beam was blocked between the fluorescence measurements. The autooxidation of 4-day old and freshly prepared solutions of DCFH was measured (Fig. 6). The increase in fluorescence, measured in counts per second (CPS) over a 40-minute period, was 1.17%/min for the fresh solution and 1.85% for the aged solution.

The hydroxyl radical generating system described above was an effective means for activating DCFH. There was over a fivefold increase in fluorescence in a sample containing

DCFH and H_2O_2-$FeSO_4$ and incubated at 37° C. At 5 min intervals, the sample was removed from the incubating bath and the fluorescence excited with the laser at 3.5 mW (Fig. 7). There was a progressive increase as the probe was increasingly oxidized. The sharp peak at 576 nm in the 5 min curve was not identified, but possibly was an emission from a metal ion contaminant (such as rubidium, based on the wavelength) in the iron sulfate reagent. This peak was never observed in the melanin-containing samples, and so was not investigated further.

The next experiment demonstrated that DCFH was activated by laser-excited melanin. A reaction mixture was made up of DCFH, bovine melanosomes, and Tris, to a total volume of 1.5 ml. This was continuously exposed to the laser excitation beam at a power of 155 mW. Very intense fluorescence was emitted with this level of excitation power. The fluorescence grew rapidly up to 20 min, and then declined by 30 min (Fig. 8). This decline was probably due to photobleaching of the probe. The fluorescence spectrum did not change over the time course of this experiment. Note that the 488 nm laser beam excited both the melanin radical and the fluorescence of the probe as it was oxidized by the melanin radicals. As will be shown below, the fluorescence of samples of DCFH and DHR123 without melanosomes increased much more slowly that of the melanosome-containing samples.

The action of light-activated melanosomes on the increase of fluorescence was dependent on the number of melanosomes added to the reaction mixture. The interpretation of the effect, however, depended on how the

Autooxidation of DCFH
RT, 1 mW excitation, 488 nm
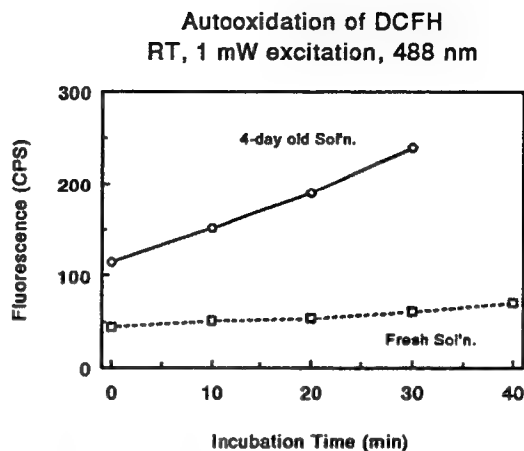
Fig. 6. Increase in fluorescence of DCFH due to oxidation. Upper curve is from a 4-day old solution; lower curve from a freshly-made solution. Excitation was with the 488 nm line from the Ar ion laser at 1 mW.

DCFH Fluorescence
Effect of hydroxyl radicals

Fig. 7. DCFH fluorescence excited by hydroxyl radicals generated by chemical reaction. Excitation of fluorescence was with the 488 nm Ar+ ion laser line at 3.5 mW.

12-11

DCFH Fluorescence
Effect of Melanin Photooxidation

File: G10
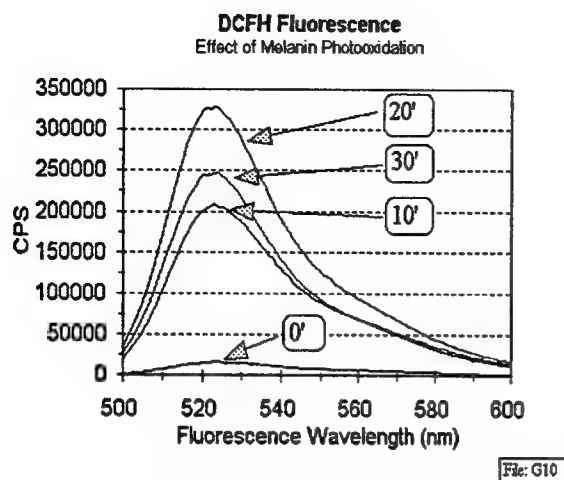
Fig. 8. Laser-activated melanosomes oxidize DCFH and produce increasing fluorescence with exposure time. Excitation was with the 488 nm Ar+ laser line at 155 mW. Note photobleaching occurred at 30' exposure.

fluorescence data was analyzed. If the raw fluorescence data in CPS were compared against the number of melanosomes in the reaction and the laser exposure time (Fig. 9A), then the peak activation of the probe appeared to occur with 62,400 melanosomes. More melanosomes seemed to quench the reactions, suggesting that self-shielding by the additional melanosomes may have occurred. On the other hand, if the relative increase in fluorescence were used, i.e. by normalizing the fluorescence readings to the initial value, then no quenching was apparent. The peak rate of fluorescence increase (but not necessarily the peak fluorescence) occurred with 156,000 melanosomes in the reaction mixture; 312,000 melanosomes produced about the same fluorescence increase. Because the fluorescence obtained with 62,400 melanosomes in the reaction gave marked increases over control regardless of whether relative or absolute fluorescence were measured, this number of melanosomes became the standard quantity in the remainder of the experiments.

The intensity of the excitation laser critically determined the magnitude of the fluores-

DCFH Absolute Fluorescence
Effect of Melanin Conc.

Files: G21-G27

DCFH Relative Fluorescence
Dependence on Melanin Conc.

Files: G21-G27
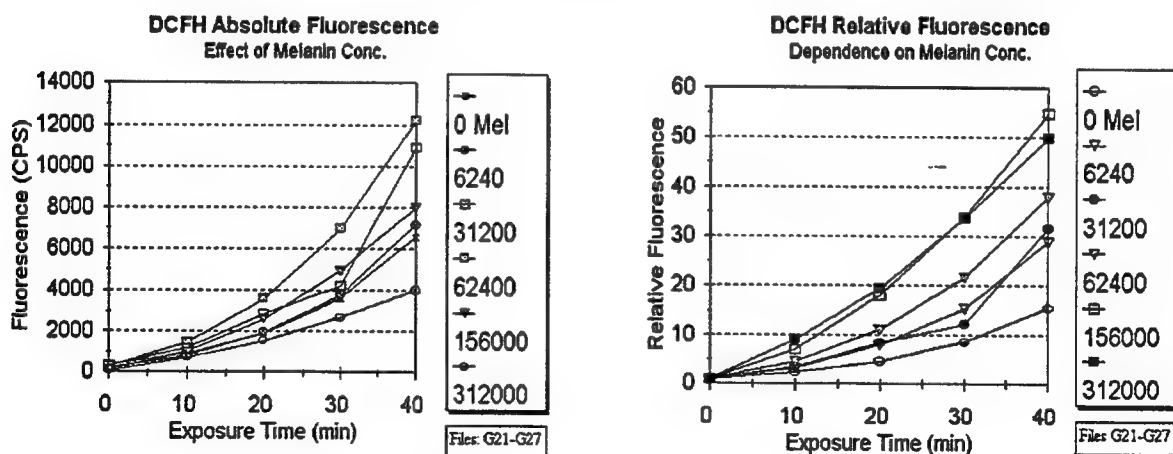
Fig. 9. Dependence of DCFH fluorescence on melanin granule density, i.e. the number of granules in the reaction mixture. A (left): DCFH fluorescence expressed as absolute number of counts/sec. B (right): DCFH fluorescence expressed relative to the CPS at the 0 time point. See text for further details.

12-12
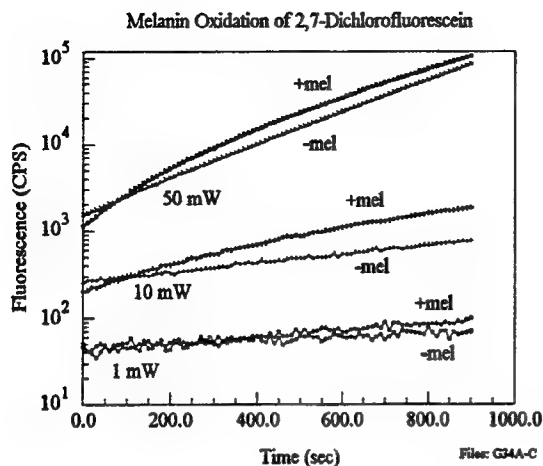
Fig. 10. Effect of excitation power on DCFH fluorescence. Reaction mixtures were excited and monitored continuously for 15 min. Fluorescence read at 522 nm. Note logarithmic axis on ordinate

cence. At very high excitation power, there was marked photobleaching, manifested by an actual decline in fluorescence. With DCFH, the effect of melanin on fluorescence increase was optimally demonstrated with an excitation power of 10 mW. Very low excitation power, e.g. 1 mW yielded low fluorescence readings, and a small but measurable difference between the samples with and without melanosomes. At higher power, e.g. 50 mW, the sample with melanosomes produced higher fluorescence than the control sample, but the relative difference declined, perhaps due to the onset of photobleaching. These data are shown for DCFH in Fig. 10. An initial attempt at modelling the kinetics of the increase in fluorescence found that a best fit function to these power curves required an expression with 3 parameters:

$$\text{Intensity} = \exp(a + b * \text{time} + c\sqrt{\text{time}}) \tag{2}$$

The form of this expression suggested a two stage mechanism might be involved in the reaction, but this has not been further investigated at this time.

To prove that the reaction between the probe and melanin involved an oxidation, the effect of ascorbic acid (AA) (vitamin C, a physiological antioxidant) was tested. In Fig. 11, the fluorescence curves from a sample of DCFH and melanosomes excited with the 50 mW laser for 15 min are shown; also in Fig. 11 are graphed the fluorescence curves produced by a similar sample with the addition of 500 μM AA, over two succesive 15 min periods. The exponential growth of fluorescence is essentially quenched by the AA and only a slow, linear increase is seen, probably due to a steady-state autooxidation of the probe. The quantitative analysis of the effect of AA on both DCFH and DHR123 will be shown in a

Fig. 11. Inhibition of DCFH fluorescence by ascorbic acid. AA present at 500 μM. Excitation with the 496 nm Ar+ laser line at 50 mW. See text for additional details.

12-13

later section.

The second probe that was tested was dihydrorhodamine 123 (DHR123). It is thought to be a lipophilic compound with a preference for membrane-bound intracellular structures such as mitochondria. Consistent with the published work with this compound[22], we found that the excitation characteristics of DHR123 were similar to that of DCFH. The peak excitation wavelength was around 500 nm, so the 496 nm Ar line was used to excite this probe. The emission spectrum in our studies peaked at 526 nm, which was a longer wavelength than the peak of the DCFH emission, but slightly shorter than the 535 nm wavelength reported

DHR123 Fluorescence
Presence or Absence of Melanin

Fig. 12. Activation of DHR123 fluorescence by laser-excited melanosomes. Excitation was with the 496 nm Ar⁺ line at 10 mW. Fluorescence was read at 526 nm at 20 (or 25) and 40 min of continuous excitation. Curves shown are the fluorescence with or without melanosomes in the reaction mixture.

for DHR123 by Royall & Ischiropoulos[22]. DHR123 fluorescence was strongly enhanced by the presence of RPE melanosomes, and the background and autofluorescence rate of DHR123 were somewhat lower that those of DCFH. Some fluorescence spectra of DHR123 are shown in Fig. 12, taken at 20-25 min and 40 min, in the presence and absence of melanosomes during continuous exposure to 496 nm light at 10 mW. There was more than a tenfold enhancement in fluorescence after 40 min of laser exposure in the presence of melanosomes (Fig. 12).

The effect of excitation power was examined over the range 1 mW to 50 mW. Samples of DHR123 with and without melanosomes were exposed continuously and the resulting fluorescence measured. As shown in Figure 13, the greatest differential was obtained with 10 mW excitation. At 1 mW excitation, the sample with melanin initially produced less fluorescence than a sample without melanin, but the fluorescence of the melanin-containing sample rapidly grew so that by 15 min it was slightly higher than the con-

Melanin Oxidation of Dihydrorhodamine 123

Fig. 13. Effect of laser excitation power on DHR123 fluorescence. Reaction mixtures were excited and monitored continuously for 15 min. Fluorescence read at 526 nm. Note logarithmic axis on the ordinate.

trol. The samples excited by 10 mW showed the same initial behavior, but by 15 min the melanin-containing sample produced nearly 10 times as much fluorescence as the control. As was the case for DCFH, with very high excitation power, i.e. 50 mW, photobleaching of the DHR123 was evident.

Effect of Antioxidants on Melanin-Induced Fluorescence. As shown in Fig. 11, photoactivated fluorescence from the probe DCFH could be inhibited by addition of an excess of ascorbic acid (AA). This inhibitory effect was examined quantitatively by determining the dose response of AA inhibition of DCFH (Fig. 14) and DHR123 (Fig. 15). In the case of DCFH, the fluorescence elicited from a mixture of the probe and melanosomes during laser excitation was very sensitive to the presence of AA; at a concentration of 1 μM, AA reduced the fluorescence below the level attained from autooxidation of the probe alone (Fig. 14A, compare the curves labelled "-Mel, 0AA" and "1μM AA"). At a concentration of 5 μM, AA reduced the fluorescence by 75%. The dose-response data for AA and DCFH is shown in Fig. 14B. From a sigmoidal function fit to the data, the ED_{50} for inhibition of DCFH fluorescence by AA was calculated to
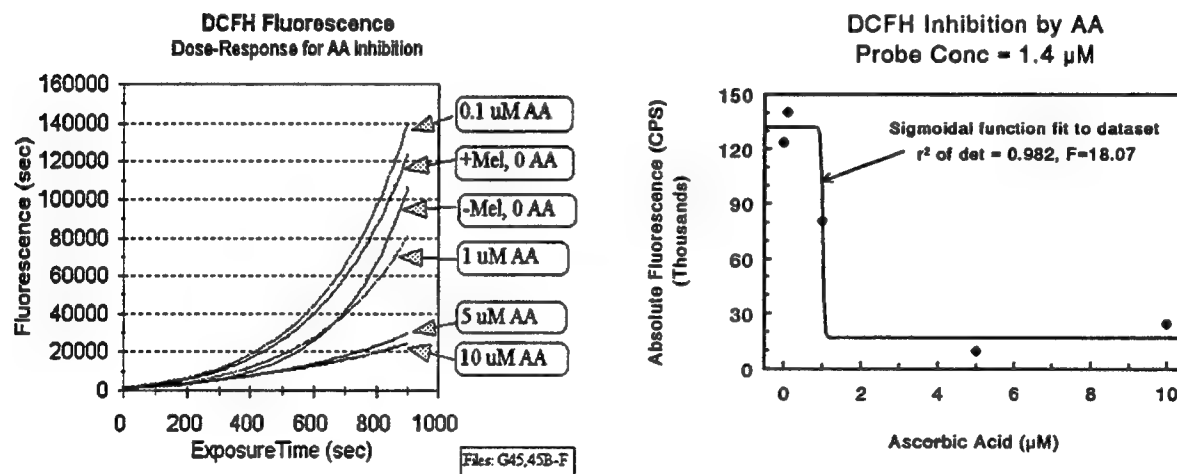
Fig. 14. Ascorbic Acid inhibition of DCFH fluorescence depends on concentration. A (left): DCFH fluorescence in CPS is shown in a family of curves for various conditions: with and without melanin, and with various concentrations of AA. Excitation power was 50 mW at 496 nm. Fluorescence was effectively quenched by AA. B (right): The dose response relation of AA and DCFH fluorescence. Equation for regresssion line is y=16958 + 114974/(1 + exp(-(x-1.0052)/-0.0241)). Solving for 50% reduction in the value of "y" (fluorescence) yielded an ED_{50} of 1.012μM for AA.

be 1.012 μM.

DHR123 was not as sensitive to the presence of AA. The curves shown in Fig 15A indicate that even at 100 μM, i.e. 20 times the probe concentration, the fluorescence was reduced by less than 40%. The dose-response curve plotted in Fig. 15B bears this out; the

12-15

DHR123 Fluorescence
Inhibition by Ascorbic Acid

DHR123 Inhibition by AA
Probe conc = 5 µM
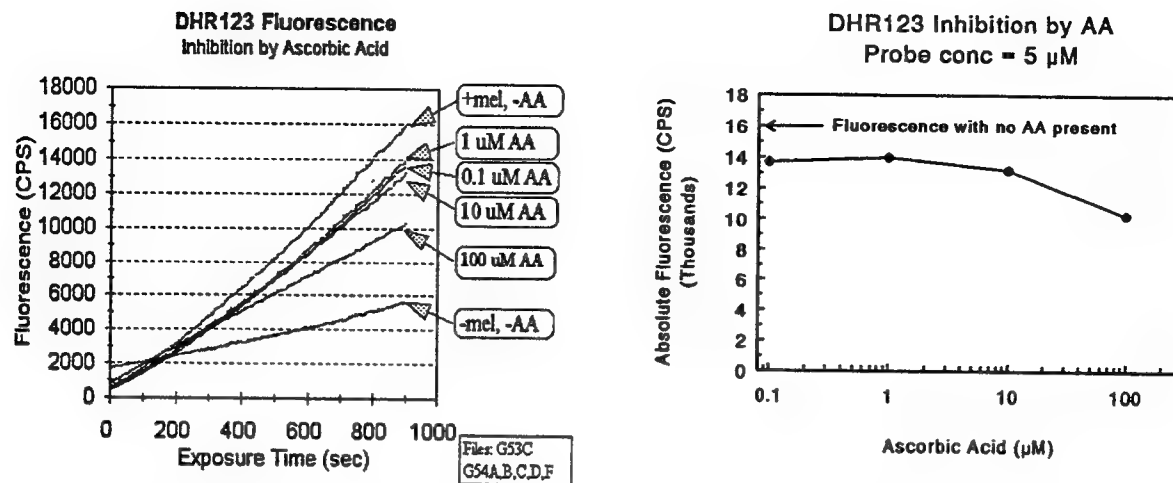
Fig. 15. Effect of ascorbic acid on DHR123 fluorescence. A left): A family of curves showing DHR123 fluorescence in the presence or absence of melanosomes, and with or without AA at various concentrations. Excitation power was 10 W at 496 nm. DHR123 fluorescence was only slightly quenched by AA. B (right): The dose-response relation of AA and DHR123 fluorescence. When AA was present at great excess (~20 fold), fluorescence was reduced, but not to the 50% level. An ED_{50} was not determined.

concentration range tested for AA was insufficient to calculate the ED_{50}.

Fluorescence Changes Induced by Native and YAG-Laser Disrupted Melanosomes. In order to anticipate the application of the fluorescence probes to *in vivo* systems, the final series of experiments in this summer project were carried out to test the probes' sensitivity to the differences in photoreactivity of native (intact) and YAG-laser disrupted bovine melanosomes. Melanosomes were disrupted by exposure to 1800 pulses from the DCR-11 Nd:YAG laser, and then their interaction with DCFH or DHR123 was excited by the 10 mW, 496 nm line from the Ar ion laser. Fluorescence from the mixtures was measured for 15 min (Fig. 16). It is clear from the increased fluorescence that disrupted melanosomes were more efficient than were the native melanosomes at oxidizing either probe (Fig. 16). It is interesting that, while the absolute levels of fluorescence from the DHR123 was greater than from DCFH, the relative increase produced by the disrupted vs the native melanosomes was about the same for both probes.

Finally, the dependence of fluorescence on the extent of melanosome disruption was examined by increasing the irradiance of the melanosome sample by the YAG laser and determining the resulting fluorescence over 15 min. As shown in Fig. 17, there was a graded, but nonlinear, increase in DCFH fluorescence as the energy per pulse was increased from 10 to 30 mJ/pulse (nominally 884 to 2,653 mJ/cm², estimating the YAG laser beam diameter as 1.2 mm). Above 30 mJ/pulse (2653 mJ/cm²), no further increase was observed, indicating that maximum disruption was achieved at that irradiance. These data are very consistent with the disruption

Effect of Nd:YAG Photodisruption
DHR123 and DCFH Fluorescence

Effect of Increasing Nd:YAG Energy
DCFH Fluoresence
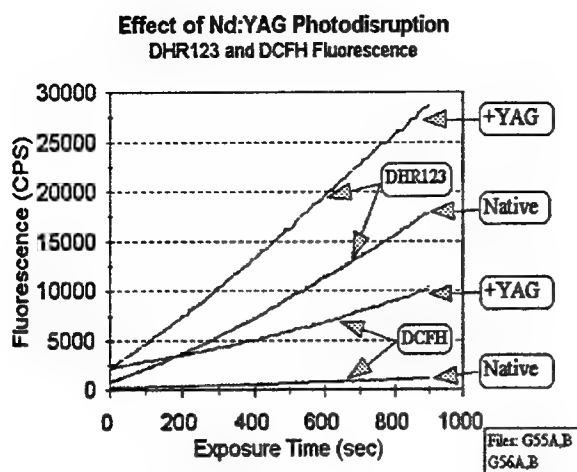
Fig. 16. Effect of YAG laser disruption on melanosome activated fluorescence of the probes DCFH and DHR123. A comparison of the native and photodisrupted melanosomes indicated that disruption increased the photoreactivity towards the probes. Excitation was 10 mW at 496 nm.
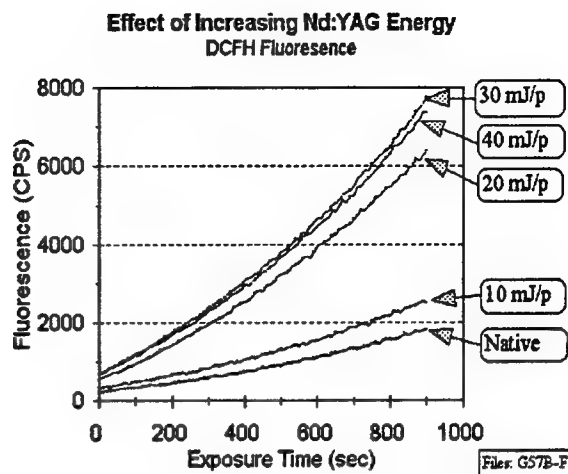
Fig. 17. Effect of YAG laser on melanosome's photoreactivity towards DCFH is graded with Q-switched laser's intensity. As the irradiance of the melanosome sample was increased, the photoreactivity of the melanosomes also icreased up to a maximum. See text for additonal details.

threshold of 500-800 mJ/cm^2 extrapolated from NADPH oxidation. Thus, in principle, a fluorescence-based assay could be used to determine melanosome granule disruption *in vivo*.

Discussion

The experiments performed during the course of the Summer Faculty Research Program served to complete some basic studies on the optical and photochemical properties of RPE melanin, and to lay the foundation for a new technique for studying tissue photosensitizers *in situ*. The data on NADPH oxidation by disrupted melanosomes will be used to complete the determination of the optical absorption coefficient for ocular melanin at a visible (532 nm) and NIR (1064 nm) wavelength[18,20]. The higher fluence required for disruption at 1064 nm is consistent with the lower absorption coefficient expected for melanin in the IR.

The increased photochemical reactivity of melanosomes disrupted by visible laser pulses has now been extended to NIR laser exposures. This aspect of melanin-light interaction probably needs to be considered when determining damage thresholds in pigmented tissue[17,19]. In terms of understanding melanin-mediated photochemical tissue reactions, the spectral content of the ambient light will also need to be considered with respect to its potential interaction with the action spectrum of the melanin radical. Although the absorption spectrum of melanin is high in the UV and declines rather smoothly into the near IR[2], the finding of a peak in the wavelength

sensitivity of the melanin radical in the range of 450 to 500 nm is new. The decline in quantum efficiency in the wavelengths below 450 nm might result from self screening due to the high absorption coefficient in the melanin sample (JA Zuclich, personal communication), however, the action spectrum data are consistent with measurements of protein photooxidation obtained by Glickman (unpublished data, but see Fig. 4 here), i.e. when melanin was present in retinal protein preparations, 488 nm light was more effective in promoting protein oxidation than was broad band light (including VIS and UVA) from a Xenon arc lamp at equal sample irradiances. The peak efficiency of the short-wavelength visible light in exciting the melanin radical may also be responsible for the so-called "blue light" retinal hazard[23].

The use of fluorescent probes to detect oxidative stress in cells has been demonstrated in some cell cultures subject to metabolic stress, e.g. vascular endothelial cells[22], but the ability to couple these probes to photoactivated melanin is a novel finding. The overall results of the current project indicate that both DCFH and DHR123 can be oxidized by the melanin radical. The appearance of the melanin-induced fluorescence is qualitatively similar to fluorescence induced by chemically produced hydroxyl radicals[21], suggesting that the probe may be oxidized either by reactive oxygen species produced by transfer of electrons from excited melanin to molecular oxygen, a mechanism previously demonstrated[8,9], or by direct interaction with an excited melanin triplet, a mechanism suggested to explain the anaerobic oxidation of AA by melanin[13].

The high efficiency of fluorescence of both probes is encouraging evidence that a usable optical signal could be obtained from the small number of probe molecules likely to penetrate single cells. A relatively small number of melanin granules in the reaction mixture is also able to activate enough probe to produce a measurable fluorescent signal. Although, according to the manufacturer, the exact oxidation-reduction potential for either probe is not known, it must be low enough to be oxidized by melanin. The redox potential for melanin (or more specifically, for the semiquinone/quinone couple likely to be responsible for the electron transfer reactions of melanin) is also not precisely known, but because light-activated melanin will oxidize ascorbic acid[12], it must be more positive than +282 mV, the redox potential for AA[24]. In this regard, it is interesting that the fluorescence of DCFH was efficiently quenched by AA at a concentration approximating the probe concentration, i.e. about 1.5 μM. This suggests that the redox potential of DCFH is higher than 282 mV, because the AA is preferentially oxidized. In contrast, DHR123 is only slightly inhibited by AA even when the concentration of AA greatly exceeds that of DHR123. One possibility for this result is that the redox potential of DHR123 is lower than +282 mV, so it is preferentially oxidized even when AA is present in excess. Other

mechanisms, however, are possible.

Regardless of the mechanism of interaction with melanin, or the inhibition by AA, the behavior of these fluorescent probes in the *in vitro* system used in this project provides a basis for further investigations. A logical progression will be to study the penetration and interactions of the probes in cultured RPE cells, and then, possibly, in animal experiments. Use of a fluorescent probe to detect increased photochemical damage *in situ* resulting from melanin granule damage following short pulse laser exposures is but one promising application for this technique.

Acknowledgements

References

1. Boettner EA, Wolter JR. Transmission of the ocular media. Invest Ophthalmol Vis Sci 1962;1:776-83.

2. Wolbarsht ML, Walsh AW, George G. Melanin, a unique biological absorber. Appl Opt 1981;20:2184-6.

3. McGinness J, Proctor P. The importance of the fact that melanin is black. J Theoret Biol 1973;39:677-8.

4. Sarna T, Pilas B, Land EJ, Truscott TG. Interaction of radicals from water radiolysis with melanin. Biochim Biophys Acta 1986;833:162-7.

5. Mason HS, Ingram DJE, Allen B. The free radical property of melanins. Arch Biochem Biophys 1960;86:225-30.

6. Cope FW, Sever RJ, Polis BD. Reversible free radical generation in the melanin granules of the eye by visible light. Arch Biochem 1963;100:171-7.

7. Riley PA. Radicals in melanin biochemistry. Ann NY Acad Sci 1988;551:111-20.

8. Sarna T, Sealy RC. Free radicals from eumelanin: quantum yields and wavelength dependence. Arch Biochem Biophys 1984;232:574-8.

9. Korytowski W, Pilas B, Sarna T, Kalyanaraman B. Photoinduced generation of hydrogen peroxide and hydroxyl radicals in melanins. Photochem Photobiol 1987;45:185-90.

10. van Woert MH. Reduced nicotinamide-adenine oxidation by melanin: Inhibition by phenothiazines.

Proc Soc Exp Biol Med 1968;129:165-71.

11. Menon IA, Haberman HF. Mechanisms of action of melanin in photosensitized reactions. In: Klaus SN, ed. Pigment Cell. Basel: Karger, 1979:345-51.

12. Glickman RD, Lam K-W. Oxidation of ascorbic acid as an indicator of photooxidative stress in the eye. Photochem Photobiol 1992;55:191-6.

13. Glickman RD, Sowell R, Lam K-W. Kinetic properties of light-dependent ascorbic acid oxidation by melanin. Free Rad Biol Med 1993;15:453-7.

14. Thompson CR, Gerstman BS, Jacques SL, Rogers ME. Melanin granule model for laser-induced thermal damage in the retina. Bull Math Biol 1996;58:513-53.

15. Toth CA, Narayan DG, Osborne C, et al. Histopathology of ultrashort laser pulse retinal damage. In: Jacques SL, ed. Laser-Tissue Interaction VII. v. 2681. Bellingham, WA: Proc. SPIE, 1996:375-81.

16. Roach WP, DiCarlo CD, Noojin GD, et al. Sub-nanosecond, single laser pulse minimum visible lesion studies in the near infrared. In: Jacques SL, ed. Laser-Tissue Interaction VII. v. 2681. Bellingham, WA: Proc. SPIE, 1996:366-74.

17. Cain CP, Toth CA, DiCarlo CD, et al. Visible retinal lesions from ultrashort laser pulses in the primate eye. Invest Ophthalmol Vis Sci 1995;36:879-88.

18. Glickman RD, Jacques SL, Schwartz JA, Rodriguez T, Lam K-W, Buhr G. Photodisruption increases the free radical reactivity of melanosomes isolated from retinal pigment epithelium. In: Jacques SL, ed. Laser-Tissue Interaction VII. v. 2681. Bellingham, WA: Proc. SPIE, 1996:460-7.

19. Glickman RD. Is there a photochemical component to retinal damage produced by ultrashort pulse laser exposures? 4th AFOSR Ultrashort Laser Pulse Collaborative Workshop, San Antonio, TX, 24-26 March, 1996.

20. Jacques SL, Glickman RD, Schwartz JA. Internal absorption coefficient and threshold for pulsed laser disruption of melanosomes isolated from retinal pigment epithelium. In: Jacques SL, ed. Laser-Tissue Interaction VII. v. 2681. Bellingham, WA: Proc. SPIE, 1996:468-77.

21. LeBel CP, Ischiropoulos H, Bondy SC. Evaluation of the probe 2',7'-dichlorofluorescin as an indicator of reactive oxygen species formation and oxidative stress. Chem Res Toxicol 1992;5:227-31.

22. Royall JA, Ischiropoulos H. Evaluation of 2',7'-dichlorofluorescin and dihydrorhodamine 123 as fluorescent probes for intracellular H_2O_2 in Cultured Endothelial Cells. Arch Biochem Biophys 1993;302:348-55.

23. Ham Jr WT, Mueller HA, Ruffolo Jr JJ, Clarke AM. Sensitivity of the retina to radiation damage as a function of wavelength. Photochem Photobiol 1979;29:735-43.

24. Buettner GR. The pecking order of free radicals and antioxidants: lipid peroxidation, α-tocopherol,and ascorbate. Arch Biochem Biophys 1993;300:535-43.

THE EFFECT OF SHORT DURATION RESPIRATORY MUSCULATURE TRAINING ON TACTICAL AIR COMBAT MANEUVER ENDURANCE AND RECOVERY

Ellen L. Glickman-Weiss
Associate Professor
School of Exercise, Leisure and Sport

Kent State University
163E MACC Annex
Kent, OH 44242-0001

THE EFFECT OF SHORT DURATION RESPIRATORY MUSCULATURE TRAINING ON TACTICAL AIR COMBAT MANEUVER ENDURANCE AND RECOVERY

Ellen L. Glickman-Weiss
Associate Professor
School of Exercise, Leisure and Sport
Kent State University

Abstract

The purpose of this study is to test the effect of a low frequency, short duration respiratory musculature training (RMT) program on tactical air combat maneuver (TACM) endurance and recovery in experienced acceleration research subjects. Also of interest is the effect of RMT on recovery from the TACM. While studies such as these have been conducted examining general strength training of the skeletal musculature, only minimal acceleration research has been conducted utilizing specific training of the respiratory musculature, despite the critical role the respiratory musculature plays in the intra-thoracic pressure component of the AGSM. Subjects will be randomly assigned to either the respiratory musculature training group (RMT) or the control group. Variables [units] to be measured include: peak heart rate response to TACM [beats·min^{-1}], TACM endurance [s], subject perceived effort (ordinal scale) during TACM, subject recovery time from TACM [s difference between TACM replications], aerobic capacity (VO$_2$ peak) [ml·kg^{-1}·min.$^{-1}$], peak and mean anaerobic power [W], miscellaneous body composition variables (fat mass, fat-free mass, etc.), maximal inspiratory/expiratory pressures [mm Hg], maximal inspiratory/expiratory capacities [ml], maximal inspiratory/expiratory flow rates [ml·s^{-1}]. While the proposed investigation underwent revisions during my 10-week tour to comprise its present content, my role was to organize the laboratory to be able to initiate data collection on this project.

THE EFFECT OF SHORT DURATION RESPIRATORY MUSCULATURE TRAINING ON TACTICAL AIR COMBAT MANEUVER ENDURANCE AND RECOVERY

Ellen L. Glickman-Weiss

Introduction

The anti-G straining maneuver (AGSM) currently taught to USAF aircrew of high performance aircraft has two principal components (1). The first, a skeletal muscular component, requires the sustained, isometric recruitment of the skeletal musculature. Second, the intra-thoracic pressure component of the AGSM involves repeated 3-s cycles of forceful expiratory efforts against a closed or partially closed glottis (1). Repetition of an effective AGSM is of critical importance to fighter aircrew in the sustained high $+G_z$ environment of air combat maneuvering to aid venous return and maintain eye-level arterial blood pressure.

Because of the extreme physical demands placed on individuals during repeated performance of the AGSM during high $+G_z$ conditions, one might assume that physically untrained individuals will fatigue sooner than their physically trained peers (2). Addressing the skeletal musculature component of the AGSM, research has been conducted to confirm and quantify the protective effect of general strength training programs (i.e., whole body resistance training) on $+G_z$ endurance (3-7). While studies such as these have been conducted examining general strength training of the skeletal musculature, only minimal acceleration research has been conducted utilizing specific training of the respiratory musculature, despite the critical role the respiratory musculature plays in the intra-thoracic pressure component of the AGSM. The research which has been conducted in acceleration utilizing respiratory musculature training (RMT) has demonstrated that such training can result in an improved AGSM (8,9) but has not been specifically applied to a simulated operational $+G_z$ endurance profile. It has been demonstrated that increased respiratory work during the AGSM results in respiratory muscle fatigue and this fatigue correlates with early termination of centrifuge exposures (10). In addition, it is possible to elicit fatigue in the diaphragm (11) and other respiratory muscles (12), and it has been demonstrated that appropriate training increases both strength and endurance of the respiratory musculature (13). However, the strength/endurance training regimens detailed in the literature are time consuming (long duration) and not practical for aircrew use (13).

The purpose of this study is to test the effect of a low frequency, short duration RMT program on the TACM endurance in experienced acceleration research subjects. Also of interest is

the effect of RMT on recovery from the TACM. This study employs the specificity of exercise training principle (14) relative to the intra-thoracic pressure component of the AGSM.

Objective

The purpose of this experiment is to evaluate the effectiveness of strength/endurance training of the respiratory musculature (short duration) in enhancing tactical air combat maneuver (TACM) endurance and recovery.

Impact Statement

This experiment is an important step in responding to the flying community's request for a training program which enhances $+G_z$ endurance. If this study is not conducted, a unique opportunity to effectively and efficiently enhance $+G_z$ endurance may be lost.

Methodology

A. Equipment and Facilities:

1. Centrifuge: Subjects' TACM endurance will be assessed from standard measurement equipment during exposures on the Armstrong Laboratory Centrifuge. Subjects will be experienced members of the acceleration subject panel and will be fitted with the Advanced Technology Anti-G Suit (ATAGS), HGU-55/P helmet and MBU-20/P oxygen mask for all centrifuge runs with the exception of a gradual onset run (GOR) to establish relaxed $+G_z$ tolerance, where no helmet/mask configuration will be worn. To minimize breathing resistance from the oxygen mask during the AGSM, a high-flow CRU-94 oxygen regulator, in the "emergency" setting, will be used for all centrifuge exposures requiring subjects to wear a mask. All subjects will be fitted with a 5-lead ECG for two-axis recording of heart rhythm and heart rate during all centrifuge exposures. Analog data signals of physiologic and anti-G suit pressure variables will be recorded on a multiple channel strip chart recorder (Gould 2800S) with magnetic tape back-up for subsequent data reduction and statistical analysis. Centrifuge subjects will be observed via closed circuit video monitors in the centrifuge control room.

2. Ergometry: Ergometry tests will be conducted in the Applied Exercise and Work Physiology Laboratory (AL/CFT) in Building 170. The Applied Exercise and Work Physiology Laboratory (AL/CFT) conducts all exercise testing according to guidelines and recommendations set forth by the American College of Sports Medicine (15).

a. Aerobic Capacity: A motorized treadmill (Quinton Q-65) will be used to elicit work for the aerobic capacity test. Expired gas will be collected from subjects during tests and analyzed utilizing a Rayfield Gas Collection System equipped with Ametek O_2/CO_2 analyzers. Blood

13-4

pressure will be monitored using a Finapres finger cuff and monitor. Subjects will also be fitted with a 3-lead ECG configuration for monitoring of heart rate and rhythm using a Quinton Q710 electrocardiograph.

b. Anaerobic Power: Anaerobic power will be determined from administration of a 30-s maximal pedaling test on a cycle ergometer (Cardionics).

3. Body Composition: Body composition measurements will be conducted in the Auxiliary Exercise Physiology and Work Physiology Laboratory in Building 170.

a. Body Density: Body density will be estimated via hydrostatic weighing using a 300-gallon hydrostatic chamber equipped with a 15 kg X 20 g autopsy scale (Chatillon). Residual volume will be determined on land prior to submersion via oxygen dilution using O_2 and CO_2 analyzers (Ametek).

b. Total Body Water: Total body water/extracellular water will be estimated utilizing a bioelectrical impedance analysis system (Xitron 4000B).

c. Anthropometry: A series of anthropometrical measurements will be made on subjects to aid in anti-G suit fitting and to serve as a secondary evaluation of body composition. Measurements include skinfold (Lange calipers), body circumference (Lufkin tape), bone breadths (Lafayette calipers), body mass (Detecto scales) and body stature (Detecto measuring rod).

4. Respiratory Variables: Respiratory tests will be conducted in the Applied Exercise and Work Physiology Laboratory (AL/CFT) and Auxiliary Exercise Physiology and Work Physiology Laboratory in Building 170.

a. Respiratory Pressures: An electronic pressure manometer (Omega) will be used to measure maximal expiratory pressure and maximal inspiratory pressure of all subjects.

b. Lung Capacities: A Collins 120-L spirometer will be used to measure lung capacities.

c. Pulmonary Flow Rates: A Collins 13.5-L spirometer will be used to measure pulmonary flow rates.

B. Subjects: Both male and female subjects from the Armstrong Laboratory acceleration subject panel who have maintained acceleration training standards will be recruited for this study. We will recruit sufficient subjects to have approximately 20 subjects (10 treatment, 10 control) complete the study. To qualify for the study, subjects must have recent $+G_z$ exposure experience and be able to complete at least four consecutive cycles (+52 s) of the TACM while wearing ATAGS. One TACM cycle consists of 5 s at $+9 G_z$, 1 s at $+5 G_z$, 5 s at $+8 G_z$ and 2 s at $+4 G_z$. Acceleration panel subjects who have not routinely performed TACM exposures with helmet and mask will perform TACM familiarization exposures wearing the full equipment ensemble, including helmet and mask, prior to establishment of maximal baseline TACM endurance times. Panel subjects also will receive RMT program indoctrination (experimental group only) before they perform respiratory testing and initiate their 6-week RMT regimen. Only acceleration panel

subjects who have no history of prolonged resistance training participation and have not participated in muscular strength training programs during the past four weeks will be studied. To control for the potentially detrimental effects of excessive aerobic training on +Gz tolerance (3), subjects who perform regular aerobic activities equal or equivalent to running more than 15 miles per week also will not be included in the study. In addition, only non-smoking subjects will be recruited.

C. Duration of the Study: Determination of relaxed $+G_z$ tolerance, acceleration variables, respiratory variables, ergometric variables and body composition variables will require approximately two weeks. These determinations will be performed prior to (pre) and following (post) the 6-week RMT period. Duration of the entire study is dependent upon the timeliness of subject recruitment. However, depending on the time required to complete pre- and post-RMT testing, duration of participation for each individual subject will be approximately 10 weeks. Every attempt will be made to have subjects adhere to the following 10 week testing/training schedule:

a. Week 1, Monday: indoctrination to respiratory testing and RMT
   (treatment group only) procedures

b. Week 1, Tuesday: centrifuge testing for $+G_z$ tolerance (pre-RMT 1)

c. Week 1, Wednesday: aerobic capacity testing (pre-RMT)

d. Week 1, Thursday: body composition testing (pre-RMT)

e. Week 1, Friday: anaerobic power testing (pre-RMT)

f. Week 2, Monday: indoctrination to respiratory testing and RMT
(treatment group only) procedures

g. Week 2, Tuesday: centrifuge testing for $+G_z$ tolerance (pre-RMT 2)

h. Week 2, Wednesday: respiratory testing (pre-RMT)

i. Weeks 3-8: RMT (5 days per week) for treatment group only and centrifuge
$+G_z$ proficiency maintenance (1 day per week) for all subjects (i.e., RMT
and control groups)

j. Week 9, Monday: re-acquaintance with respiratory testing procedures

k. Week 9, Tuesday: centrifuge testing for $+G_z$ tolerance (post-RMT 1)

l. Week 9, Wednesday: aerobic capacity testing (post-RMT)

m. Week 9, Thursday: body composition testing (post-RMT)

n. Week 9, Friday: anaerobic power testing (post-RMT)

o. Week 10, Monday: re-acquaintance with respiratory testing procedures

p. Week 10, Tuesday: centrifuge testing for $+G_z$ tolerance (post-RMT 2)

q. Week 10, Wednesday: respiratory testing (post-RMT)

D. Experimental Procedures:

　　1. General Experimental Design: The general experimental design entails a single between-subjects effect and no within-subject effects. There are two levels of the between-subjects effect (i.e., RMT, control). Subjects will be randomly assigned to either the RMT or the control group. Variables [units] to be measured include:

> a. peak heart rate response to TACM [beats·min.$^{-1}$]
>
> b. TACM endurance [s]
>
> c. subject perceived effort (ordinal scale) during TACM
>
> d. subject recovery time from TACM [s difference between TACM replications]
>
> e. aerobic capacity (VO$_2$ peak) [ml ·kg^{-1}·min.$^{-1}$]
>
> f. peak and mean anaerobic power [W]
>
> g. miscellaneous body composition variables (fat mass, fat-free mass, etc.)
>
> h. maximal inspiratory/expiratory pressures [mm Hg]
>
> i. maximal inspiratory/expiratory capacities [ml]
>
> j. maximal inspiratory/expiratory flow rates [ml·s^{-1}]

　　2. Centrifuge Procedures:

　　　　a. Testing for +G$_z$ Tolerance: +G$_z$ tolerance will be tested for pre- and post-RMT training. Although peak relaxed +G$_z$ tolerance will be established and periodically re-assessed for all subjects, the principal criterion by which effectiveness of the exercise conditioning regimens will be evaluated is the time (s) subjects can endure the TACM profiles. For all centrifuge runs, the centrifuge will be configured with the upright (13° seat back angle) ACES II-type seat. Foot plates in the centrifuge gondola will be positioned to provide an approximately 120° knee joint angle. Each subject will be properly fitted with ATAGS and a HGU-55/P helmet with MBU-20/P mask for all centrifuge exposures with the exception of the GOR to establish relaxed +G$_z$ tolerance where no helmet/mask will be worn. The pre- and post-RMT centrifuge testing will be conducted in duplicate. Acceleration profiles for pre- and post-RMT data collection consist of:

> i. GOR (0.1 +G$_z$·s^{-1}) to peak +G$_z$ or other endpoint criteria (medical monitor decision, self-initiated decision, nausea, 100% peripheral light loss or 50% central light loss*), relaxed (no AGSM), anti-G suit worn but NOT pressurized
>
> ii. Warm-up consisting of 1 cycle of ROR (6.0 +G$_z$·s^{-1}) to +5.0 G$_z$ for 15 s, with AGSM, with anti-G suit pressurized
>
> iii. 2-min. rest period
>
> iv. TACM to fatigue or other end-point criteria (medical monitor decision, self-initiated decision, nausea; 100% peripheral light loss or 50%

central light loss*), with AGSM, with anti-G suit pressurized, no central observer coaching

v. 5-min. rest period

vi. TACM to fatigue or other end-point criteria (medical monitor decision, self-initiated decision, nausea, 100% peripheral light loss or 50% central light loss*), with AGSM, with anti-G suit pressurized, no central observer coaching

*if $+G_z$ loss of consciousness (G-LOC) occurs it will constitute termination of $+G_z$ endurance testing and, consequently, time to G-LOC (s) will be used as the endurance variable.

b. $+G_z$ Proficiency Maintenance: In order to maintain their entry level AGSM proficiency and $+G_z$ endurance baseline, subjects will be required to perform a TACM centrifuge training profile once a week throughout the RMT period. This $+G_z$ training requirement consists of a warm-up chosen by the subject, followed by a submaximal TACM profile (i.e., ≈75% of the pre-RMT TACM performance) during first four weeks of RMT. During weeks 5 and 6 of the RMT, subjects will be allowed to perform maximally during their weekly centrifuge "maintenance" runs (i.e., subjects will be allowed to perform the TACM to fatigue). Those subjects in the control group, who will not be engaging in RMT, will be required to follow the same centrifuge maintenance schedule as those subjects participating in RMT. Any subjects missing more than one of their weekly centrifuge maintenance rides over the 6-week training period will be removed from the study.

3. Ergometry Procedures: Tests for aerobic capacity and anaerobic power will be conducted pre- and post-RMT training.

a. Testing for Aerobic Capacity: Standard electrocardiographic electrodes will be applied prior to performing the test for aerobic capacity so heart rhythm and rate can be monitored. Blood pressure will be monitored throughout the test for aerobic capacity via oscillometry using the Finapres monitoring system. Expiratory gasses will be collected and analyzed using the Rayfield Gas Analysis System. The test will continue to volitional or physiological maximal effort, which ever occurs first. Specific reasons for test termination include (15,16):

i. subject request to terminate test

ii. subject is unable to continue test (i.e., volitional maximum)

iii. failure of subject oxygen consumption to increase with an increase in work demand with confirmation from elevated respiratory exchange ratio (≥1.15) and/or heart rate greater than age estimated maximum

iv. onset of angina or angina-like symptoms

v. significant drop (20 mm Hg) in systolic blood pressure or failure of systolic blood pressure to increase in response to an increase in work demand

vi. excessive rise in blood pressure (systolic>260 mm Hg; diastolic>115 mm Hg)

vii. subject exhibits signs of poor perfusion such as lightheadedness, confusion, ataxia, pallor, cyanosis, nausea or cold/clammy skin

viii. ECG abnormalities such as a pronounced ECG change from baseline (i.e., > 2mm of horizontal/down sloping ST-segment depression or >2mm of ST-segment elevation), less serious arrhythmias (supraventricular tachycardia) and exercise-induced bundle branch block indistinguishable from ventricular tachycardia

ix. dyspnea

x. equipment failure

The treadmill protocol for aerobic capacity is that derived by Ellestad (17). This protocol requires subjects to exercise continuously to VO_{2peak} according to the following grade/belt rate progression:

Stage	Minutes	Grade (%)	Belt rate ($mi \cdot h^{-1}$)
1	0:00-3:00	10.0	1.7
2	3:01-5:00	10.0	3.0
3	5:01-7:00	10.0	4.0
4	7:01-10:00	10.0	5.0
5	10:01-12:00	15.0	5.0
6	12:01-14:00	15.0	6.0
7	14:01-16:00	15.0	7.0
8	16:01-18:00	15.0	8.0

b. Testing for Anaerobic Power: The Wingate Test (18,19) will be used to determine peak and mean anaerobic power. The test is a 30-s maximal cycle ergometer test. The warm-up for the test consists of 1) 1 min. of pedaling (50 rpm) against no resistance; 2) 3 min. of pedaling (50 rpm) against a moderate load (50% of actual test resistance) interspersed with three sprints (maximal rpm) of 5 s each and 3) 1 min. of pedaling (50 rpm) against no resistance. After a brief rest period (3-5 min.), the subject pedals as fast as possible against the test workload (0.09 kg·kg

body mass^{-1}) for 30 s. Recovery includes pedaling against a minimal resistance ($\approx$1.0 kg) for at least 3-5 min. Premature termination of the test can occur for the following reasons:

     i. subject request to terminate test

     ii. subject is unable to continue test

     iii. onset of angina or angina-like symptoms

     iv. subject exhibits signs of poor perfusion such as lightheadedness, confusion, ataxia, pallor, cyanosis, nausea or cold/clammy skin

     v. dyspnea

     vi. equipment failure

4. Body Composition Procedures: Body composition measurements will be made pre- and post-RMT training. Subjects will be asked to report for testing in a post-absorptive state following an overnight fast. Subjects will be asked to maintain normal hydration during this fasting period.

a. Body Density: Body volume will be determined via hydrostatic weighing. From this body volume estimate body density will be calculated for use in a multi-component model with total body water to estimate total body adiposity. Subjects will be weighed underwater while forcefully expiring to residual lung volume. The procedure will be repeated until successive attempts fail to produce an increase in underwater weight. Residual lung volume will be measured on land prior to the hydrostatic weighing procedure. The modified oxygen dilution method (20) for measuring residual lung volume will be used. This technique requires subjects to forcefully expire to residual volume, hyperventilate 100% O_2 and forcefully expire to residual volume again. The gas collected after the final expiration will be analyzed using Ametek O_2 and CO_2 analyzers.

b. Total Body Water: Bioelectrical impedance analysis will be used to estimate total body water content. Total body water will be used in a multi-component model with body density to estimate total body adiposity. The procedure requires subjects to lie supine and have four superficial electrodes placed on the right side of his/her body at the following anatomical sites:

     i. posterior aspect of the hand over the distal portion of the metacarpals

     ii. posterior aspect of the arm just superior to the carpals

     iii. superior aspect of the foot over the distal portion of the metatarsals

     iv. anterior aspect of the leg just superior to the tarsals

A measurement current of 500μA is passed between the electrodes for the determination of the impedance to this electrical flow provided by the circuit (i.e., the subject's body). This device is listed with the U.S. Food and Drug Association as a Class II Medical Device and has been designed to meet the safe current limits standards for patient connected equipment. The impedance data will be reduced using the Xitron 4000B with computer interface.

c. Anthropometry: Subjects will undergo a series of anthropometrical measurements for body composition estimation (21, 22) and to assist in anti-G suit fit. The anthropometrical measurements include skinfolds, body circumferences, bone breadths, body mass and stature.

5. Procedures for Respiratory Musculature Testing: Measurements of inspiratory and expiratory maximal pressures, capacities and flow rates will be made pre- and post-RMT training. Testing will include:

a. Maximal Inspiratory Pressure: The determination of maximal inspiratory pressure requires subjects (seated with nose clip applied) to first expire maximally to residual lung volume through a plastic cylinder connected to a pressure manometer (Omega) which will measure negative pressure during the subsequent inspiration. When maximal expiration has occurred the subject will inspire maximally. Shortly after inspiration begins ($\approx$0.5 s), the cylinder will be occluded requiring the subject to inspire against a closed circuit. The inspiration will last approximately 3 s and a reading of negative pressure made. Subjects will be instructed how to perform the task to ensure minimal pressure production by the facial muscles. The test will be conducted in triplicate.

b. Maximal Expiratory Pressure: The determination of maximal expiratory pressure requires subjects (seated with nose clip applied) to first inspire maximally to total lung capacity through a plastic cylinder connected to a pressure manometer (Omega) which will measure positive pressure during the subsequent expiration. When maximal inspiration has occurred the subject will expire maximally. Shortly after expiration begins ($\approx$0.5 s), the cylinder will be occluded requiring the subject to expire against a closed circuit. The expiration will last approximately 3 s and a reading of positive pressure made. Subjects will be instructed how to perform the task to ensure minimal pressure production by the facial muscles. The test will be conducted in triplicate.

c. Inspiratory Capacity: The determination of inspiratory capacity requires subjects (seated with nose clip applied), at the end of resting expiration, to inspire maximally to total lung capacity through a plastic cylinder connected to a spirometer (Collins) to measure gas volume. The test will be conducted in triplicate.

d. Expiratory Capacity: The determination of expiratory capacity requires subjects (seated with nose clip applied), at the end of resting inspiration, to expire maximally to residual lung volume through a plastic cylinder connected to a spirometer (Collins) to measure gas volume. The test will be conducted in triplicate.

e. Maximal Inspiratory Flow Rate: The determination of maximal inspiratory flow rate requires subjects (seated with nose clip applied) to first expire to residual volume and then inspire maximally through a plastic cylinder connected to a spirometer (Collins). Flow/volume curves will

be recorded such that flow rates can be calculated at various points throughout the inspiration. The test will be conducted in triplicate.

f. Maximal Expiratory Flow Rate: The determination of maximal expiratory flow rate requires subjects (seated with nose clip applied) to first inspire to total lung volume and then expire maximally through a plastic cylinder connected to a spirometer (Collins). Flow/volume curves will be recorded such that flow rates can be calculated at various points throughout the expiration. The test will be conducted in triplicate.

g. Maximal Voluntary Ventilation (Maximal Breathing Capacity): The determination of maximal voluntary ventilation requires subjects (seated with nose clip applied) to breathe as deeply and rapidly as possible for 15 s through a plastic cylinder connected to a spirometer (Collins) to measure gas volume. The test will be conducted in triplicate.

6. Procedures for Respiratory Musculature Training: RMT will be conducted over a 6-week period. RMT subjects will train five days per 6-day week (i.e., training sessions can be scheduled Monday-Saturday). Within the 6-week RMT regimen, subjects will be allowed to miss a total of six training sessions. If the subjects miss more than six training sessions or miss consecutive training sessions, they will be removed from the study. A training supervisor will indoctrinate all subjects in the treatment group (i.e., not control subjects) on the RMT tasks. All RMT will require subjects to be seated in an ACES II upright seat (13° seatback angle), located in the Auxiliary Exercise Physiology Laboratory in Building 170, which is similar to the one used in the centrifuge gondola. The foot plates on this mock-up will be situated so that the subject knee joint angle is at approximately 120°. Each training session will consist of:

a. Respiratory Strength Training: RMT subjects (seated with nose clip applied) will be asked to inspire to total lung capacity with an open glottis and maintain for 10 s. Subjects will then expire maximally to residual volume and maintain for 10 s. The respiratory cycle will be repeated three times in week 1 and week 2 of the RMT program. To provide an overload, the respiratory cycle will be repeated four times in weeks 3 and 4 and will be repeated five times in the final two weeks of the RMT program. Time to complete this portion of the RMT session will be approximately 3-5 min.

b. Resistance Endurance Training: RMT subjects (seated with nose clip applied) will be asked to inspire/expire through a mouthpiece connected to a plastic tube. The tube will be narrowed to provide resistance to air flow. The stenosis will be such to provide a resistance which results in an inspiratory pressure equal to approximately 70% of the pre-RMT maximal inspiratory pressure. To provide an overload, the duration of this resistive respiration bout will be 8 min. in week 1 of the RMT program and will increase by 1 min. each week after that (i.e., on week 6 of the RMT program the duration will be 13 min.). Time required to complete this portion of the RMT session will range from 10-15 min.

13-12

8. Statistical Analysis: A sample consisting of 20 subjects (RMT group n_1=10 and control group n_2=10) is desired to enable the reasonable likelihood of detecting truly significant effects. An analysis of variance (1 treatment with 2 levels) will be utilized where the subjects' aerobic and anaerobic capacities will be introduced as covariates into the general linear model. The level of significance will be set a priori at $p \leq 0.05$.

References

1. Gillingham KK, Fosdick JP. High-G training for fighter aircrew. Aviat. Space Environ. Med. 1988; 58: 12-9.

2. Gillingham KK. High-G stress and orientational stress: physiological effects of aerial maneuvering. Aviat. Space Environ. Med. 1988; 59(11, Suppl.): A10-A20.

3. Burton RR. Simulated aerial combat maneuvering tolerance and physical conditioning: current status. Aviat. Space Environ. Med. 1986; 57: 712-714.

4. Epperson WL, Burton RR, Bernauer EM. The influence of differential physical conditioning regimens on simulated aerial combat maneuvering tolerance. Aviat. Space Environ. Med. 1982; 53(11): 1091-1097.

5. Epperson WL, Burton RR, Bernauer EM. The effectiveness of specific weight training regimens on simulated aerial combat maneuvering G-tolerance. Aviat. Space Envrion. Med. 1985; 56: 534-539.

6. Mizumoto C, Iwane M. The effects of isotonic training on $+G_z$ tolerance. Japanese J. Aerospace Environ. Med. 1984; 25: 15-29.

7. Tesch PA, Hjort H, Balldin UI. Effects of strength training on G tolerance. Aviat. Space Environ. Med. 1983; 54(8): 691-695.

8. Geng X. The strengthening of respiratory muscles and the $+G_z$ tolerance. Chinese J. Aviat. Med. 1991; 2(3): 133-136.

9. Wang S, Zhing J, Geng X. Enhancement of anti-G effect of L-1 maneuvering by strengthening of respiratory muscles in pilots. Chinese J. Aviat. Med. 1993; 4(2): 90-92.

10. Bain B, Jacobs I, Buick F. Respiratory muscle fatigue during simulated air combat maneuvering (Abstract). Aviat. Space Environ. Med. 1995; 66(5): 470 (#41).

11. Roussos CS, Macklem PT. Diaphragmatic fatigue in man. J. Appl. Physiol. 1977; 43: 189-197.

12. Roussos CS, Fixley M, Gross D Macklem PT. Fatigue of the inspiratory muscles and synergic behavior. J. Appl. Physiol. 1979; 46: 897-904.

13. Leith DE, Bradley M. Ventilatory muscle strength and endurance training. J. Appl. Physiol. 1976; 41: 508-516.

14. Sale D, MacDougall D. Specificity in strength training: a review for the coach and athlete. Science Periodical on Research and Technology in Sport. Ottawa, The Coaching Association of Canada, March 1981.

15. American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription (5th Edition). Baltimore, Williams and Wilkins. 1995: 77-78.

16. Howley ET, Bassett DR, Welch HG. Criteria for maximal oxygen uptake: Review and commentary. Med. Sci. Sports Exerc. 1995; 27: 1292-1301.

17. Ellestad MS. Stress Testing Principles and Practice (3rd Edition). Philadelphia, F.A. Davis. 1986.

18. Bar-Or O. A new anaerobic capacity test: characteristics and applications. Proceedings of the 21st World Congress in Sports medicine. Sep 7-12, 1978, Brasilia.

19. Bar-Or O. The Wingate anaerobic test: an update on methodology, reliability and validity. Sports Medicine. 1987; 4: 381-394.

20. Wilmore JH, Vodak PA, Parr RB, Girandola RN, Billing JE. Further simplification of a method for determination of residual lung volume. Med. Sci. Sports Exerc. 12; 1980: 216-218.

21. Jackson AS, Pollock ML. Generalized equations for predicting body density of men. Br. J. Clin. Nutr. 1978; 40: 497-504.

22. Jackson AS, Pollock ML, Ward A. Generalized equations for prediction body density of women. Med. Sci. Sports Exerc. 1980; 12: 175-182.

DEVELOPMENT OF A PHYSIOLOGICALLY-BASED PHARMACOKINETIC MODEL FOR THE UPTAKE OF VOLATILE CHEMICALS DURING A SHORT TIME INTERVAL: PRELIMINARY STAGES OF MODEL DEVELOPMENT

Irwin S Goldberg
Professor
Departments of Physics and Engineering

St. Mary's University
One Camino Santa Maria
San Antonio, TX 78228-8534

DEVELOPMENT OF A PHYSIOLOGICALLY-BASED PHARMACOKINETIC MODEL FOR THE UPTAKE OF VOLATILE CHEMICALS DURING A SHORT TIME INTERVAL:
PRELIMINARY STAGES OF MODEL DEVELOPMENT

Irwin S. Goldberg
Professor
Departments of Physics and Engineering
St. Mary's University

Abstract

The development of a physiologically-based pharmacokinetic model has been initiated to simulate the physiological uptake of halogenated hydrocarbon vapors from the lungs into the arterial blood and into the tissue compartments. This model will facilitate the investigation of health risks associated with the use of the halogenated hydrocarbons that are used as fire-fighting agents. Specifically, the chemical uptake during a five-minute exposure to halothane is investigated. Also the modeling is extended to include the uptake of other volatile chemicals.

The emphasis of the research involves the preliminary stages in the development of the mathematical model. The applications of the mass-balance principles used for the development of the model is introduced. The underlying assumptions are examined. A brief research literature review is presented. Criteria for characterization of the compartmental time scales are presented. Alternative non-linear models are described and evaluated. The role of the concentration effect and the second gas effect upon alveolar uptake is reviewed. The effects of the absorption of chemicals upon the surfaces of the conducting airways is discussed along with the applications of numerical fluid dynamic methods for the quantification of these effects..

DEVELOPMENT OF A PHYSIOLOGICALLY-BASED PHARMACOKINETIC MODEL FOR THE UPTAKE OF VOLATILE CHEMICALS DURING A SHORT TIME INTERVAL: PRELIMINARY STAGES OF MODEL DEVELOPMENT

Irwin S. Goldberg

I. Introduction

Halon 1301 has been commonly used as a fire-fighting agent in occupied spaces. However, because of the ozone-depletion properties of this chemical, halogenated hydrocarbons have been proposed as replacements for Halon 1301. These halogenated hydrocarbon replacement chemicals are likely to have cardiac sensitization potential at some exposure concentration, and they therefore present health risks. Cardiac sensitization is considered to be a function of the concentration of the chemical in heart tissue (Vinegar 1996). The development of a mathematical model, to simulate the physiological uptake of halogenated hydrocarbon vapors from the lungs, would facilitate the investigation of health risks associated with the use of these chemicals. The emphasis of the investigation described in this report involves the preliminary stages in the development of such a mathematical model. Specifically, I will discuss the uptake during a five-minute exposure to halothane, but I will also discuss the manner in which the model can be extended to include the uptake of other volatile chemicals.

For toxicological applications, the model should be useful for the quantitative prediction of dosage so that alterations in the dosage level, variations of the chemical agent, anatomical differences among different individuals, and physiological differences and similarities among different species can be incorporated into the model. For this purpose, physiologically-based pharmacokinetic modeling is used (Bishoff 1992, Williams et al. 1996). That is, the model is to be constructed using known anatomical details and physiological parameters with principles of physics and chemistry applied. In this manner, a quantitative predictive model is constructed based upon underlining knowledge, rather than the more empirical curve fitting approach.

A model describing the uptake and distribution of halogenated hydrocarbons consists of a system of interconnected and interacting tissue compartments. Mass-balance principles are used to develop the system of ordinary differential equations used to describe the exchange of the chemical agent between these compartments.

II. Background:

Previous investigations undertaken by the author involved a mathematical model for the simulation of the maintenance dosage of halothane during closed-circuit surgical anesthetic procedures (Goldberg et. al. 1978). That study involved the calculation of the maintenance dosage which is required after the initial four minutes of anesthetic induction. For that calculation, transient delays involving the lung uptake and the blood circulation time were not considered. Also, nonlinear effects of the anesthetic upon ventilation, cardiac output and the blood-flow distribution to tissue compartments were not considered. In contrast, effects involving pulmonary and circulatory delays and nonlinearaties should be considered in a model involving the transient behavior of a volatile drug during a dosage exposure time lasting less than five minutes.

III. The Basic Transient Model

When constructing the compartmental model, the following assumptions are made:

1. All compartments are well mixed, and the chemical concentration of the blood draining each compartment is always in equilibrium with the blood in that compartment.

2. Metabolic losses of the chemical agent are negligible.

3. Cutaneous losses of the chemical agent are negligible.

4. Lateral diffusion of the chemical agent from well-perfused organs (such as muscle) into fat is negligible.

5. There is no facilitated transport or active transport of the chemical agent.

6. Ventilation is treated as continuous.

7. Diffusion of the chemical across the alveolar membrane, from the alveolus into the capillary, is complete within the transit time of blood through the pulmonary capillary network. Therefore, the concentration of the chemical in the arterial blood leaving the lung is in equilibrium with the concentration of the chemical in the alveolar gas.

Assumption 1 (all compartments are well mixed) may lead to error in the computation of chemical uptake, especially at the early stages of uptake to slowly equilibrated or poorly perfused compartments. However, because of the very small quantities of anesthetic contained in the blood draining these slowly equilibrating compartments, this error would be expected to be small during the early stages of equilibration. Additionally, that assumption has been experimentally shown to lead to accurate predictions of washin or washout of inert gases from the entire body over time frames of minutes to hours (Wagner, 1987).

Assumption 2 (metabolic losses are negligible) is based upon the smallness of the losses of chemicals due to metabolism, especially during a five minute exposure time. Thus, the metabolic losses are not expected to be significant when calculating the direct chemical dosage and uptake. However, in studies involving metabolic products and their toxicological effects, the rates of production of metabolites would need to be considered, especially during long-time exposures. As demonstrated by Williams et. al. (Williams et. al., 1996), a compartmental model can be altered so that the kinetics involving metabolic transformations can be included.

Assumption 6 (respiration can be treated as continuous) has been shown by Riggs and Goldstein to be a legitimate simplification for lumped-parameter compartmental models involving uptake of inert gases of low-to-moderate solubility (Riggs and Goldstein, 1961). In their calculations, simulated results from compartmental models involving cyclic ventilation were compared with those involving steady ventilation. Their results show that the more extensive equations, introduced by including the effects of cyclic respiration upon compartmental gas exchange, has no practical advantage over the more simplified formulation involving steady respiratory flow. Furthermore, as explained when the transfer function and frequency characteristics are discussed later in this report, the compartments behave as low-pass filters in which the respiratory oscillations of approximately 0.25 Hz would be effectively damped out. Thus, according to a lumped-parameter compartmental model, the omission of effects related to respiratory flow oscillations appears to be further justified. However, these justifications for omitting oscillatory respiratory effects should be reviewed with caution since they are based upon lumped-

parameter compartmental models in which time-varying volume changes and details of blood flow/ventilation inequalities and inhomogenities are ignored. Also, when very highly-soluble chemicals are involved, absorption and desorption of the chemical at the epithelial tissue and mucus lining of the nose and the upper airways must be considered; in these cases, the effects of the cyclic nature of respiration must be included in the model (Fiserova-Bergerova, 1980; Medinsky et. al., 1993). These effects are discussed in section IX.

Assumption 7 (the chemical concentration of the arterial blood leaving the lung is in equilibrium with the alveolar gas) has been determined to be very accurate according to mathematical models involving diffusion of non-reacting chemicals across the alveolar membrane. According to calculations by Forster (Forster 1957), the tension of dissolved non-reactive gases in the blood of the pulmonary capillary will equal 99% of the alveolar tension in approximately 0.01 seconds of transit time: this time is much smaller than the mean total transit time of blood through the pulmonary capillary network. This hypothesis has been further supported by clinical experiments in which measured end-tidal halothane concentration was shown to closely reflect the measured arterial blood concentration (Stoelting 1973).

The model, describing the transient uptake of a volatile chemical, can be broken down into three interconnected and interacting submodels consisting of the tissue group submodel, the circulatory submodel, and the lung submodel. A block diagram describing such a model is shown in figure 1. The tissue group submodel consists of seven parallel tissue compartments. Each tissue compartment is perfused by arterial blood at a concentration equal to C_a. The circulatory submodel consists of an arterial pool, three tissue-group venous pools, a central venous pool, and a peripheral (skin) shunt. The lung submodel consists of the lung mixing compartment and the lung shunt. Further details of each of these submodels are given below. The blood-flow junctions are represented by the shaded octagons in figure 1. Flows leaving a junction by different paths are assumed to have identical concentrations. The concentration of the flow leaving each junction is calculated as the average of the concentrations of the flow paths entering that junction, weighted according to the flow rates of the entering flow paths. That is, for a junction with N entering flow paths, the concentration, C_{out}, for every flow path leaving that junction is represented symbolically as

$$
C_{out} = \frac{\sum_{j=1}^{N} C_j Q'_j}{\sum_{j=1}^{N} Q'_j} \quad , \tag{1}
$$

where C_j and Q'_j represent the concentration and flow rate of each flow path entering the junction (In this report, flow variables, are denoted with prime superscripts .)

Figure 1

U'_A ↑ ↓ U'_A

$(1-f s) Q'_T$

lung mixing — $\{C_L\}$

Q'_T $\{C_v\}$

$\{C_v\}$ $f s Q'_T$ pulmonary shunt Q'_T $\{C_a^*\}$

central venous pool $\{C_a\}$

$\{C_v^*\}$ $\{C_{vpv}\}$ venous pool viscera $\{C_{vp^*}\}$ kidney Q'_1 arterial pool

brain Q'_2

$\{C_{vpl}\}$ heart Q'_3

venous pool lean tissue $\{C_{vpl^*}\}$ liver Q'_4 Q'_5

muscle Q'_6

venous pool adipose tissue $\{C_{vpf^*}\}$ conn. tissue Q'_7 Q'_8

$\{C_A\}$ $\{C_{vpf}\}$ adipose tissue

Q'_8 peripheral skin shunt

$\{C_k\}$ = concentration; Q'_k = blood flow rate; U'_A = alveolar ventilatory flow rate

It should be noted that figure 1 represents a preliminary model in its early stage of development, and that several changes of this model may be necessary before a working model is developed. In fact, in the process of model development, successful preliminary models lead the investigator to ask questions, test hypothesis, and suggest experiments leading to the development of more accurate and sophisticated models.

IV. The Tissue Submodel

The tissue submodel consists of seven tissue compartments. As shown in figure 1, these include the kidney, brain, heart, liver, muscle, connective tissue and adipose tissue compartments (Goldberg et. al., 1978; Lerou et. al., 1991). Mass-balance principles are used to develop the differential equations describing the chemical uptake of each compartment. Consider a well-mixed tissue compartment labeled with subscript k . The resulting differential equation describing C_k, the fractional concentration of the chemical agent leaving compartment k , is

$$\frac{dC_k}{dt} = \frac{Q_k}{\lambda_k V_k}(C_a - C_k) ,$$ (2)

where Q'_k is the flow rate of blood to compartment k, λ_k is the blood/tissue solubility coefficient of the chemical contained in compartment k, V_k is the volume of compartment k, and C_a is the time-dependent chemical concentration of the arterial blood entering compartment k. The total amount of chemical contained in compartment k at any time (excluding the anesthetic contained in the blood equilibrated with compartment k) is equal to the product $\lambda_k V_k C_k$.

14-**6**

Experimental data for the solubility coefficients of halothane in various tissue compartments can be found in reports by Lowe (Lowe, 1972; 1981). Data for tissue volumes can be found in the report "Reference Man" of the International Commission on Radiological Protection (ICRP, 1975). The perfusion, Q'_k, to the various tissue compartments can be determined from the review by Cowles, Borgstedt, and Gilles (Cowles, Borgstedt, and Gilles 1971). Data for the volumes and perfusions to various tissue compartments are summarized by Davis and Mapleson (Davis and Mapleson 1981). Data for the alterations of cardiac output caused by a 2% arterial halothane concentration are given by Smith et. al. (Smith et. al., 1972).

When the parameters Q'_k, V_k and λ_k are independent of the calculated concentrations, equation 2 is linear. However, often the flow to the various compartments, Q'_k, are influenced by the local chemical concentration or by the chemical concentration in the brain or in the arterial blood; in these cases, the equations become nonlinear (Munson et. al., 1973; Fukui and Smith, 1981 a b; Aberecht et. al., 1977; Eger at. al ., 1970; Ashman et. al., 1970).

Even though nonlinear effects may need to be considered, it is instructive to look at the analytic behavior of equation 2 under hypothetical linear conditions. Assuming linearity, the following mathematical properties of the compartmental model can be determined:

1. The solution to equation 2 shows that the time-dependent relationship between $C_k(t)$ and $C_a(t)$ can be expressed by the convolution integral

$$C_k(t) = \left(\frac{Q_k{}'}{\lambda_k V_k}\right) \int_0^t C_a(t-\tau)\exp\left(\frac{-Q'_k}{\lambda_k V_k}\ \tau\right)d\tau \ , \tag{3}$$

where $C_k(t)$ denotes the concentration of the chemical in the blood draining the tissue compartment k at time t, and τ is a dummy integration variable (see appendix 1). Equation 3 shows that each compartment can be considered as an exponentially-weighted cumulative delay between the input variable C_a and the output variable C_k.

2. The parameter ($V_k\lambda_k/Q'_k$) can be considered as a generalized time constant: that is, with the specific hypothetical conditions where $C_a(t)$ is a step function, $C_k(t)$ would equilibrate to 63% of $C_a(t)$ when t is equal to ($V_k\lambda_k/Q'_k$). Thus, the parameter ($V_k\lambda_k/Q'_k$) can be thought of as a generalized response time of the compartment k: the compartment is referred to as a "fast compartment" when this parameter is small and as a "slow compartment" when this parameter is large.

3. Additionally, using a linear-systems analog, a transfer function can be calculated in terms of input C_a and output C_k for each compartment. The frequency response solutions, obtained from these transfer functions, demonstrate that each compartment can be considered to behave as a low-pass filter (see appendix 2). That is, when time oscillations of $C_a(t)$ have periods that are small compared to ($V_k \lambda_k/$

Q'_k), these oscillations are damped by the system response of compartment k and they can be neglected. In that case, only the steady flow-averaged value of concentration is calculated.

With consideration of the range of the parameters ($V_k \lambda_k / Q'_k$) involved when nonlinear effects are included , the above three mathematical properties are useful in the selection of alternative assumptions to be made, effects to consider or neglect, and approximations to be made in the development of the model.

V. The Circulatory Submodel

The circulatory submodel consists of the arterial and venous pools in addition to a peripheral circulatory shunt as shown in figure 1. The mathematical model for each pool (assumed to be well mixed) is given by the mass balance relationship

$$\frac{dC_{out(p)}}{dt} = \frac{Q'_p}{V_p}\left(C_{in(p)} - C_{out(p)}\right) ,$$
(4)

where $C_{out(p)}$ represents the concentration of the chemical in blood draining the pool p, $C_{in(p)}$ represents the concentration of the chemical in blood entering the pool p, Q'_p represents the blood flow through pool p and V_p represents the volume of the pool p. For a linear approximation, $C_{out(p)}$ is related to $C_{in(p)}$ through a convolution relationship similar to equation 3 (with λ set equal to unity). Thus, each arterial or venous pool can be considered as an approximation to an integrated cumulative delay element in the model.

In a more extensive arterial-delay model, the single arterial pool shown in figure 1 is replaced with a linearly-connected cascade of small arterial pools (Jacquez, 1983 pages 164-165; MacDonald, 1978). Variations in the circulatory delay times from the heart-lung system to the various tissue compartments can then be modeled by varying the number of arterial delay elements in the array connected to each tissue compartment (Mapleson ,1973) .

The central circulation time (from the local vein, through the lungs, to the local artery) can be estimated from Oldendorf's measurement of the mean transit time from the antecubital vein to the corotid artery as 12 seconds (Oldendorf 1962). Of this, approximately 9.5 seconds can be attributed to transit through the heart chambers and pulmonary circulation. Therefore, only approximately 2.4 seconds is left for transit through the main veins and arteries. Thus, the distances of particular tissues and organs from the heart would be unlikely to have a major influence upon the circulation times. This suggests that the modeling of the arterial delays to the various organs and tissue compartments can be accurately approximated by a single arterial pool. (Davis and Mapleson 1981). Furthermore, calculations by Mapleson (Mapleson 1973) determined that the single-arterial-pool storage model, as a means of simulating circulation times, gives results which closely approximate the more extensive simulations of circulatory delays.

The peripheral circulatory shunt flow has been estimated as approximately 15% of the total cardiac output (Lerou et. al., 1991; Davis and Mapleson, 1981). This flow is related to thermal regulation. Therefore, for

simulation of chemical exposure during fire-fighting activities, the model must allow for variations in the peripheral circulatory shunt flow under heat-stress conditions.

VI. The Lung Submodel

The lung submodel consists of the lung mixing compartment and the pulmonary shunt as shown on figure 1. This submodel describes the uptake of the chemical agent from the inspired air through the lungs into the arterial mixing pool.

The tracheobronchial tree of the lungs can be divided into the conducting airways (consisting of the trachea, the bronchi, and non-respiratory bronchioles) and the respiratory airways (consisting of the respiratory bronchioles and the alveoli) (Hlastala and Berger 1996, Horsefield 1985, West 1990, Weibel 1963). Since respiration does not take place in the conducting airways, the volume of inspired gas that remains in the conducting airways at the end of inspiration is referred to as the anatomical dead space of the lungs.

Additionally, the effect of the volume of air going to the unperfused alveoli and to underperfused alveoli acts as dead space ventilation, known as alveolar dead space. The sum of the anatomical dead space and the alveolar dead space is known as the physiological dead space. The volume of the anatomical dead space in a normal adult is approximately in the range of 150 to 180 ml. In a young, normal individual, the volume of the physiological dead space is only slightly greater, about 25% to 35% of the average tidal volume.

The time constants (equal to the volume/flow rate ratio) for the mixing of gas in the dead space compartment of the lung are less than 0.05 minutes. Since the volume of the dead space compartment is small (indicating very small uptake) and time constants are short (indicating fast response time), the differential equation involving mixing in the dead space does not need to be considered, at least for the approximation in which effects of oscillatory ventilation are neglected. Using mass-flow conservation principles, the alveolar ventilation rate, U'_A, is calculated as

$$ U'_A = \left(\frac{V_T - V_D}{V_D} \right) U'_E \ , \tag{5} $$

where V_T denotes tidal volume (approximately 0.5 liters), U'_E denotes inspired ventilatory flow rate (approximately 7 liters/minute), and V_D denotes dead space volume (approximately 0.15 liters).

As shown in figure 1, a pulmonary shunt fraction , f_S , denotes the fraction of the total cardiac output, Q'_T, which bypasses the ventilatory activity of the lung. The remainder of the cardiac output, equal to $(1-f_S)Q'_T$, perfuses the lung mixing compartment where the exchange process takes place between the pulmonary blood and the alveolar gasses. In young, healthy individuals the shunt fraction, f_S , is commonly less than 5% and can usually be neglected in calculations of chemical uptake.

As shown in figure 1, the lung mixing compartment is perfused with mixed-venous blood at a chemical concentration equal to C_v and a flow rate equal to $(1-f_s)Q'_T$. Also that compartment is ventilated with gas of a chemical concentration equal to the inspired chemical concentration, C_E , and with an alveolar air flow rate equal

to U'_A. Using mass balance principles, the differential equation for the chemical concentration, C_A, in the alveolar gas of the lung mixing compartment is

$$\frac{dC_A}{dt} = \frac{U'_A}{V_{LG} + \lambda_B V_L}\left(C_E - C_A\right) + \frac{\left(1 - f_s\right)Q'_T\left(C_V - \lambda_B C_A\right)}{V_{LG} + \lambda_B V_L} \quad , \tag{6}$$

where λ_B is the blood/gas partition solubility coefficient ($\lambda_B = 2.3$ for halothane), V_{LG} is the "alveolar space" lung gas volume, and V_L is the volume of alveolar tissue and blood in which the anesthetic agent is distributed (where the solubility coefficient in lung tissue is approximately equal to that in blood).

The chemical concentration of the blood leaving the lung mixing compartment is C_L (see figure 1), where

$$C_L = \lambda_B C_A. \tag{7}$$

Also, C_a^* is determined from the instantaneous mixing of C_L with C_V (see figure 1) as

$$C_a^* = (1 - f_s)\,C_L + f_s\,C_V \tag{8}$$

The concentration, C_a, of the blood leaving the arterial pool and perfusing each of the seven tissue compartments is determined by

$$\frac{dC_a}{dt} = \frac{Q'_T}{V_{ap}}\left(C_a^* - C_a\right) \quad , \tag{9}$$

where V_{ap} is the volume of the arterial pool.

VII. The Concentration Effect

In the derivation of equations 5 and 6, we neglected the loss of the volume of gas from the lung due to absorption of the gas into the blood. In the mass-balance relations, this volume loss can be neglected when the fractional concentration of the inhaled gas is low. However, when a chemical with appreciable solubility and at a high alveolar fractional concentration is transported across the alveolar membrane, equations 5 and 6 must be modified to account for the "concentration effect" (Eger, 1963; Eger, 1964). Under these conditions, a large fraction of the gas within the alveoli is transported across the alveolar membrane into the blood with each respiratory cycle. The flow rate of the gas exhaled from the alveoli would therefore be significantly smaller than the flow rate of the gas inhaled into the alveoli. The mass balance relation for the volume flow rate of the alveoli becomes (Munson et.al., 1973)

$$U'_{A(IN)} = U'_{A(EX)} + W' \quad , \tag{10}$$

where $U'_{A(EX)}$ denotes the flow rate of gas exhaled from the alveoli, $U'_{A(IN)}$ denotes the flow rate of gas inhaled into the alveoli, and W' denotes the transport rate of the chemical across the alveolar membrane. (see figure 2)

Note the contrast between the result of equation 10 and the result from the assumption made when the concentration effect was neglected in section VI; in that approximation, $U'_{A(EX)}$ and $U'_{A(IN)}$ were assumed to be equal, with both denoted simply by U'_A as determined by equation 5.

The accumulation of the chemical in the alveolar lung gas is given by

$$V_{L(Tot\,Gas)}\,\frac{dC_A}{dt} = C_E U'_{A(IN)} - C_A U'_{A(EX)} - W' \,, \tag{11}$$

Where $V_{L(Tot\,Gas)}$ is the total lung gas volume (approximately 2.7 liters) and C_E is the inspired concentration.

Figure 2
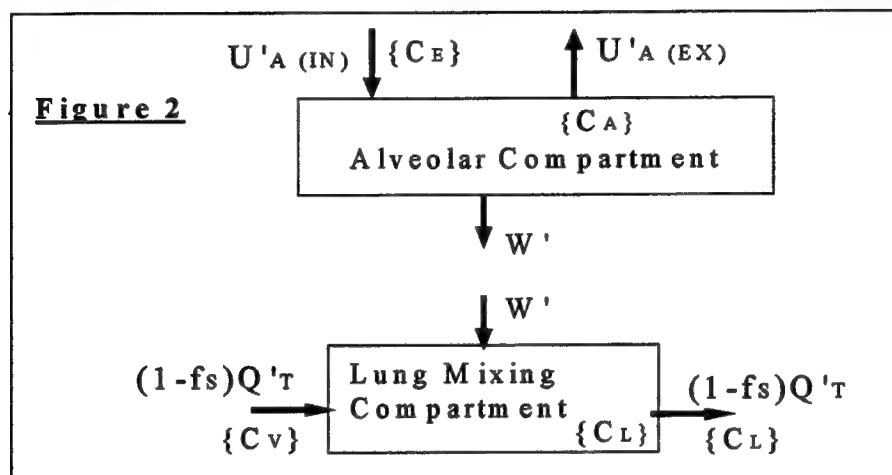
The concentration, C_L, of the chemical in the blood draining the lung is determined by the equation

$$V_L \frac{dC_L}{dt} = (1 - f_s)\,Q'_T\,(C_V - C_L) + W' \,, \tag{12}$$

where $C_L = \lambda_B\,C_A$ as in equation 7. Combining equations 10, 11, and 12 with equation 7, we can eliminate $U_{A(IN)}'$ and W', with the result (Munson et.al., 1973)

$$\left[V_{L(Tot\,Gas)} + (1 - C_E)\lambda_B V_L\right]\frac{dC_L}{dt} = \tag{13}$$
$$U'_{A(EX)}(\lambda_B\,C_E - C_L) + (1 - f_s)(1 - C_E)\lambda_B Q'_T(C_V - C_L)$$

Equation 13 can be combined with equations 8 and 9 of section VI to determine C_a. When $C_{A(IN)} \ll 1$, the concentration effect can be neglected: in that case, equation 13 reduces to equations 6 and 7.

The concentration effect increases when the total tissue uptake of the chemical increases. Since uptake is related to solubility, the concentration effect increases with chemicals of higher solubility.

In applications for clinical anesthesia, halothane gas is usually utilized at a low concentration of about 2%. Therefore, the concentration effect is usually not included in the models of halothane uptake often reported in the clinical research literature. However, this effect needs to be included in the models for toxicological investigation when higher concentrations of chemicals may be involved. Also, when two or more chemicals are

inhaled, the concentration effect from a volatile chemical at a high concentration may affect the uptake of the other chemicals; this effect is known as the "second gas effect" (Fahri and Olszowka 1968).

VIII. Nonlinear Effects

Total cardiac output (Q_T'), alveolar flow rate (U_A'), and distribution of blood flow to the various tissue groups (Q_k') may be affected by the local concentration of the chemical within the tissue or by the concentration of the chemical in the brain or in the arterial blood. Concentration-dependent changes in the ventilatory flow or in the compartmental circulatory flow would affect the rate of uptake of the chemical agent and would therefore cause the differential equations to become nonlinear. In these cases, analytic solutions to these differential equations become very difficult to obtain, and numerical methods, such as Runge-Kutta-Fehlberg method, must be used (Forsythe et. al., 1977; Press et. al., 1986). Several available computer software packages, such as MATLAB (Math Works Inc. Natick, Mass.) or ACSL (MGA Software, Concord, Mass.) , provide interactive routines for constructing the model simulations, obtaining the numerical solutions of the nonlinear systems of differential equations, and producing graphical presentations of the results.

In the following three subsections, different nonlinear models will be described. Since these models have their limitations as discussed below, more work remains to be done so that these models can be incorporated into a comprehensive model to be used for toxicological studies.

VIII. 1. Nonlinear model of Munson , et. al.

Nonlinear effects of halothane upon the total cardiac output and upon respiration were investigated by Munson , et. al. (Munson , et. al. 1973). In their investigation, they used empirical response curves to model nonlinear effects of halothane concentration in the brain upon the alveolar ventilation flow rate and effects of halothane concentration in arterial blood upon total cardiac output. Their nonlinear model had the following limitations: (1) As pointed out by the authors, dose-dependent alterations in the distribution of the cardiac output between the various tissue groups were not considered. (2) Because the intended application of the simulations involved clinical anesthesiology, the empirical data supporting the model was limited to maximum halothane concentrations of 2%..

The simulated results from their model demonstrate that, as the concentration of chemical in the brain increases, the resulting reduced ventilation will inhibit the further uptake of the chemical. However, due to a delay between a rise in the arterial concentration and a responsive rise of the concentration in the brain, the model predicts a transient overshoot in the arterial chemical level during a 0-5 minute initial transient interval.

VIII. 2. Nonlinear model of Smith , et. al.

Smith et al. (Smith et al. 1972) developed a nonlinear model in which the effects of halothane concentration upon the cardiac output and blood flow distribution were considered. Their nonlinear model had the following limitations: (1) Possible nonlinear effects of halothane concentration upon alveolar ventilation was neglected. That is, alveolar ventilation was considered as a constant 4 liters /min. (2) Their assumed linear relation between cardiac output and halothane concentration appears doubtful (Fukui and Smith 1981, Eger et al 1970). (3)

14-**12**

Also, they assumed that circulatory data during steady state conditions can be applied to circumstances where the concentration of halothane is continuously changing.

In this model, cardiac output and the blood flow distribution to various compartments was altered by nonlinear modifications of either the total cardiac output or the distribution of the flow conductances of each compartment. That is, a parallel-resistance electrical circuit-flow analog was used to model the flow distribution for N tissue compartments.

Their model involved twelve compartments in which two brain compartments, consisting of a gray-matter brain compartment and a white-matter brain compartment, were considered. Alternative models can be evaluated; that is, cardiac output and the changes in the conductance can be controlled by halothane concentration in either the arterial blood, the brain (gray matter) or the myocardium.

Their results does not appear to show a large deviation from the linear model during the first five minutes of dosage during clinical anesthetic delivery. However, for our purposes, nonlinear effects need to be evaluated for the higher range of halothane concentrations that may be encountered in toxicological investigations. Also, a more comprehensive model needs to be constructed in which the nonlinear effects of flow redistribution examined in the model by Smith et al is combined with the nonlinear ventilation effects examined in the model by Munson et. al. These comprehensive models need to be supported by additional experiential data.

VIII. 3. Nonlinear model of Fukui and Smith .

A very extensive nonlinear model was developed by Fukui and Smith (Fukui and Smith 1981a) which involves 88 equations and 124 parameters and which incorporates three distinct but interdependent submodels. These submodels include: (1) a tissue uptake model consisting of 18 compartments, (2) a breath-by breath pulmonary model, consisting of two compartments which describes the air pressure-flow relations in the airways, and (3) a beat-by beat cardiovascular model, consisting of 15 compartments, which describes in detail the blood pressure-flow relations.

Alterations in the regional vascular resistance (or reciprocal compliance) were modeled to vary as a function of halothane concentration in arterial blood. Additionally, a baroreceptor feedback loop is included in which an increase in arterial blood pressure amplitude causes the heart rate to decrease. Specifically, the assumed relationship between the heart rate and the mean arterial blood pressure is described by

$$[HR] = \frac{1}{K_1[AP] + K_2} , \qquad (14)$$

where [HR] is the heart rate, [AP] is the mean arterial pressure, K_1 is the linear baroreceptor constant, and K_2 is the set-point baroreceptor constant. To model the effect of halothane concentration upon the baroreceptor characteristics, both K_1 and K_2 were assumed to decrease as a linear function of the halothane concentration in the brain (gray matter). Also, the amplitude of the reciprocal of the heart compliance, which is related to cardiac

contractility and stroke volume, can be programmed to decrease as a linear function of the halothane concentration in the arterial blood, the gray matter of the brain and/or the heart. The cardiac output, Q'_T, is then calculated as the product of the heart rate and stroke volume. As acknowledged by Fukui and Smith, their simulated results showed that pulsate events during cardiac and respiratory cycles have little or no effect upon halothane uptake or distribution.

In a second paper, Fukui and Smith added a CO_2 control loop to the model (Fukui and Smith 1981b). The production of CO_2 was assumed to take place in seven tissue compartments, and a 15-compartment mass-transport model was added to the basic model to simulate the transport and production of CO_2. The model included effects of the partial pressure of CO_2 upon ventilation, myocardial contractility (the reciprocal of left ventricular compliance), the baroreceptor-heart-rate control loop and regional vascular resistance. Also, the manner in which increasing concentration of halothane attenuates these CO_2 induced responses was investigated. In the model, the mathematical relationships describing the effects of CO_2 and holothane upon ventilation, myocardial contractility, and the baroreceptor-heart-rate control loop were assumed to be linear. Effects that may be incorporated into this model include: (1) the modulation of ventilatiory flow rate by the concentration of CO_2 in the brain, (2) the modulation of myocardial contractility and regional vascular resistance (conductance) by halothane concentration in various selected tissue compartments, (3) the modulation of the blood-flow conductance to various tissue compartments by the concentration of CO_2, (4) the modulation of the effect of the CO_2 responses by the concentration of halothane, and (5) the influence of both halothane and CO_2 concentrations upon the slope and set point constants for the baroreceptor pressure- heart-rate feedback control loop. These effects of CO_2 upon chemical uptake become especially significant for toxicological investigations involving fire-fighting conditions.

IX. The Role of the Upper Respiratory Tract in the Uptake of Volatile Chemicals

For the lung submodel, described in section VI, uptake of the chemical agent was assumed to take place exclusively in the alveolar region of the lung. The functional role of the conductive airways of the upper respiratory tract in the uptake of chemicals has been ignored. Also, the cyclic nature of the flow was assumed to be unimportant. These assumptions are invalid and they cause the calculated uptake to be overestimated when highly soluble chemicals are considered (Fiserova-Bergerova, 1980; Medinsky et. al., 1993).

The lower than expected respiratory uptake of highly volatile polar solvents has been suggested to be due to the following phenomena (Medinsky, 1993; Johanson, 1991): During inhalation, a polar solvent (with a very large blood solubility) is partially absorbed by, or dissolves in, the surface of the respiratory epithelium. Likewise, during expiration, desorption of some of the solvent occurs from the respiratory epithelium into the expired gas stream. As a result, the amount of solvent vapor that reaches the alveoli is reduced, and the relative respiratory uptake is thereby reduced. This absorption/desorption process is a consequence of not only blood solubility of the chemical agent but also the cyclic nature of the gas exchange process (Landahl and Herrmann, 1950; Schrikker et. al., 1985).

The effect involving the cyclic absorption/desorption process described above becomes important when the blood/air solubility coefficient becomes sufficiency large such that a significant fraction of the inhaled chemical is absorbed at the lining of the conducting airways during a time interval corresponding to the inspiratory phase of the breathing cycle. Thus, this effect does not appear to be important for halothane and other anesthetic agents which commonly have a blood/gas solubility coefficient less than ten. However, this effect is important for chemicals such as ethanol, methanol, or methoxyethinol which have blood/gas solubility coefficients exceeding several hundred.

Current research for the investigation of the deposition of highly volatile chemicals on the upper airway surfaces includes (1) multicompartmental models of mass transfer between the airway lumen, the lung tissue, and the capillary blood network at cascaded sections of the conducting airways, (2) Investigation of the effects of the solubility, reactivity and diffusitivity of the chemical upon absorption and transverse transport across the airway surface (Hanna et.al.,1989), and (3) flow visualization and the development of mathematical models of convective transport and diffusion of gases in the nose and in the conductive airways (Medinsky et.al, 1993). For this purpose, models have been constructed to describe regional transport of inhaled gases by flow convection coupled with diffusion within the airways using equations describing conservation of mass and momentum. The equation describing diffusion with convection is given as (Bird et.al ., 1960; Watson, 1983; Scherer et. al., 1972)

$$\frac{\partial C}{\partial t} + \overline{U} \bullet \nabla C - D\nabla^2 C = 0 \quad, \tag{15}$$

where C is the concentration of the chemical, t denotes time, ∇ denotes the gradient operator, and ∇^2 denotes the Laplacian operator. The velocity vector, $\overline{U}$, in equation 15 is determined from the solution to the Navier-Stokes and continuity equations describing fluid motion. The flow models are complicated because of the turbulent mixing in the larger airways, secondary motion and mixing due to three-dimensional entrance flow behavior near the bifurcations, the cyclic nature of respiratory flow, the elastic properties of the airways, the curvature of the airway tubes, the variations in the lengths of alternate flow paths of the bifurcating flow streams, and the coupled transport effects of convection, streaming, and diffusion. (Pedley et.al, 1977).

For developing flow near the entrance of tube, an entrance length can be defined as the distance from the entrance at which the flow profile is within 1% of the developed-flow solution. Analytical results show that, even for slow axisymmetric flow at very small Reynolds numbers[1], the entrance length of straight tubes exceeds one tube radius (Goldberg and Folk , 1988; Goldberg et. al. 1990). Additional analytical results show that, for non-axisymmetric entrance flow at very small Reynolds numbers, the entrance length exceeds two tube radii (Zou

---

[1] The Reynolds number is a dimensionless parameter. For flow in a circular tube of diameter d , the Reynolds number is defined as Re= U'd/ν, where U' is the average velocity and ν is the kinematic viscosity of the fluid. Flow calculations at very small velocities are referred to as Stokes flow approximations. For these approximations, the Reynolds numbers are negligibly small, and the Navier-Stokes equations of fluid motion become approximately linear.

et. al., 1993). These analytical estimates of entrance lengths are based upon the calculated rates of decay of the exponentially-decaying modes of the eigenfunction expansions for the linearized solutions for time-dependent Stokes flow at very slow flow velocities in straight, rigid tubes. In contrast to these approximations, the flow in the conducting airways involves the effects of (1) large Reynolds numbers, (2) curved tubes, and (3) non-rigid tubes: each of these effects would be expected to cause the entrance length to increase. Thus the entrance lengths near the bifurcations of the conducting airways would be expected to exceed two tube radii, even in the smallest airways. Since the conductive airways consists of a bifurcating assembly of short tubes, many with length/diameter ratios less than four (Weibel, 1963), a large part of the behavior of the convective process involves entrance flow phenomena. Thus, to determine the amount of absorption of chemical on the airway surfaces, we must solve the Navier-Stokes equations involving entrance flow near the bifurcations coupled with equation 15 describing convective diffusion. Rigorous numerical methods, such as finite-element methods (Carey and Oden, 1986) or control-volume methods (Patankar, 1980) are needed for these calculations.

Appendix 1.   Derivation of equation 3

Taking the Laplace transform of equation 2, and collecting terms we obtain

$$C_k^L(s) = \left(\frac{Q_k'}{\lambda_k V_k} \right) \left(\frac{1}{S + \frac{Q_k'}{\lambda_k V_k}} \right) C_a^L(S) \ , \tag{A1-1}$$

where the superscript L denotes Laplace-transformed function, and S denotes the transform variable (In the derivation of equation A1-1, all initial conditions are assumed to equal zero.) The term

$$\left(\frac{1}{S + \frac{Q_k'}{\lambda_k V_k}} \right)$$ is equal to the Laplace transform of $\exp\left(\frac{-Q_k' t}{\lambda_k V_k} \right)$ STEP(t) , where STEP(t) is a unit step function..

The inverse Laplace transform of a product of two Laplace-transformed functions is equal to the convolution of the two time-domain inverse-transformed functions (Kreyszig 1988, Kaplan, 1981). Therefore, the convolution result given by equation (3) follows from the inverse transform of equation (A1-1). Note that the step-function behavior is incorporated into the zero lower limit of integration of the convolution integral of equation 3.

Appendix 2.   Frequency-Response Characteristics of the Compartmental Mixing Model

Taking Fourier transforms of the dependent variables of equation 2, we obtain

$$i\omega\, C_k^F(\omega) \;=\; \frac{Q_k'}{\lambda_k V_k}\Big(C_a^F(\omega) - C_k^F(\omega)\Big),$$

(A2-1)

where the superscript F denotes Fourier-transformed function, ω denotes the transform variable, and $i = \sqrt{(-1)}$. Considering Ca as the input to the system and C_k the output, a complex transfer function is defined as

$$T(\omega) \equiv \frac{C_k^F(\omega)}{C_a^F(\omega)}.$$

(A2-2)

Rearranging terms of equation A2-1 and dividing, the complex transfer function is obtained as

$$T(\omega) = \frac{Q_k'}{\lambda_k V_k}\, \frac{1}{\left(i\omega + \dfrac{Q_k'}{\lambda_k V_k}\right)},$$

(A2-3)

The magnitude of the transfer function is

$$|T(\omega)| \;=\; \frac{\left|C_k^F(\omega)\right|}{\left|C_a^F(\omega)\right|} \;=\; \frac{Q_k'}{\lambda_k V_k}\, \frac{1}{\sqrt{\omega^2 + \left(\dfrac{Q_k'}{\lambda_{ki} V_k}\right)^2}},$$

(A2-4)

where | | denotes magnitude or complex absolute value. With a lumped-parameter, linear system approximation, these results have the following interpretation: If the input, Ca, oscillates in time with frequency ω, the output, C_k, oscillates at the same frequency, and the ratio of the amplitudes, $|C_k| / |Ca|$, is equal to $|T(\omega)|$ given in expression A2-4. Expression A2-4 demonstrates a low-pass frequency filter characteristic. That is, oscillations in the concentration leaving a compartment k is damped out if the frequency of the oscillation, ω, is sufficiently large such that $\omega \gg Q_k'/(\lambda_k V_k)$, (or equivalently, if the period of the oscillation is much smaller than $(\lambda_k V_k / Q_k')$. Specifically, using a signal analysis analogy, the response characteristics for each compartment k is that of a first-order low-pass filter with a corner frequency equal to $Q_k'/(\lambda_k V_k)$ and with a 6dB/octave asymptotic rolloff at higher frequencies.

References

Albrecht, R.F., Miletich, D.J., Rosenberg, R., and Zahed, B., *Cerebral Blood Flow and Metabolic Changes from Induction to Onset of Anesthesia with Halothane or Pentobarbital,* Anes., vol 38, pp. 252-256, 1977.

Ashman, M.N., Blesser, W.B., and Epstein, R.M., *A Nonlinear Model for Uptake and Distribution of Halothane in Man,* Anes., vol 33, pp. 419-429, 1970.

Bird, R.B. Stewart, W.E., and Lightfoot, E..N., *Transport Phenemona*, John Wiley and Sons, New York, N.Y., 1960.

Bischoff, K.B., *PBPK Models: What Are We Really Assuming?*, CIIT Activities, vol. 12, pp. 1-6, 1992.

Carey,G.F., and Oden, J.T., *Finite Elements in Fluid Mechanics*, Prentice-Hall, Englewood Cliff, N.J., 1986.

Cowles, A.L., Borgstedt, H.H., and Gilles, A.J., *Tissue Weights and rates of Blood Flow in Man for the Prediction of Anethetic Uptake and Distribution,* Anes., vol. 35, pp. 523-526, 1971.

Davis, N.R., and Mapleson, W.W., *Structure and Quantification of a Physiological Model of the Distribution of Injected Agents and Inhaled Anesthetics,* Brit. J.Anaesth., Vol 53, pp. 399-405, 1981.

Du Bois, A.B. , and Rogers, R.M., *Respiratory Factors Determing the Tissue Concentrations of Inhaled Toxic Substances*, Reapir. Physiol. vol. 5, pp 34-52, 1968.

Eger, E.L, *Effects of Inspired Anesthetic Concentration on the Rate of Rise of Alveolar Concentration,* Anes., vol. 24, pp. 153-157, 1963.

Eger, E.L, *Respiratory and Circulatory Factors in Uptake and Distribution of Volatile Anesthetic Agents,* Brit. J. Anaesth, vol.36, pp. 155-171, 1964.

Eger, E.L, Smith, N.T., Stoelting, R.K., Cullen, D.J., Kadis, L.B., and Whitcher, C..E., *Cardiovascular Effects of Halothane in Man,* Anes., vol 32, pp. 396-409, 1970.

Farhi, L.E., and Olszowka, A.J., *Alalysis of Alveolar Gas Exchange in the Presence of Soluble Inert Gases,* Respir. Physiol., vol 5, pp.53-67, 1968.

Fiserova-Bergerova, V., *Physiological Models for Pulmonary Administration and Elimaination of Inert Vapors and Gases,* in Modeling of Inhalation Exposure to Vapors: Uptake, Distribution and Elimination, Vol. 1, Fiserova-Bergerova, V., ed., Chapter 4, pp.73-100, 1980.

Forsythe, G.E., Malcolm, M.A., and Molar, C.B., *Computer Methods for Mathematical Computations,* Prentice Hall, 1977.

Foster, R.E., *Exchange of Gasses Between Alveolar Air and Pulmonary Capillary Blood; Pulmonary Diffusing Capacity,* Physiol. Rev., vol 37, 1957.

Fukui, Y., and Smith, N.T., *Interactions Among Ventilation, The Circulation, and the Uptake of Halothane, Use of a Hybrid-Computer Multiple Model, I. The Basic Model,* Anes., vol 54, pp.107-118, 1981(a).

Fukui, Y., and Smith, N.T., *Interactions Among Ventilation, The Circulation, and the Uptake of Halothane, Use of a Hybrid-Computer Multiple Model, II Spontaneous vs. controlled ventilation, and the effects of CO_2 ,* Anes., vol 54, pp. 119-124, 1981(b).

Goldberg, I.S., Carey, G.F., McLay, R., and Phinney, L., *Periodic Viscous Flows; A Benchmark Problem,* Intl. Journ. for Numerical Methods in Fluids, vol 2, pp. 187-197, 1990.

Goldberg, I.S., and Folk, R.T. *Solutions for Steady and Nonsteady Entrance Flow in a Semi-Infinite Circular Tube at Very Low Reynolds Numbers,* SIAM Journ Appl Math, vol. 48, pp.770-791, 1988.

Goldberg, I.S., Mostert, J.W., Lanzl, E.F., and Lowe, H.J., *A Pharmacokinetic model of Closed-Circuit Anesthesia,* Annals of Biomedical Engineering, vol. 6, pp. 231-249, 1978.

Guyton, A.C., *Basic Human Physiology: Normal Function and the Mechanisms of Disease,* W.B Sanders Co., Philadelphia, Pa.., 1971.

Hanna, L.M., Frank, R.F., and Sherer, P.W., *Absorptrion of Soluble Gases and Vapors in the Respiratory System,* Vol. 40, Lung Biology in Health and Disease, (C. Lenfant, ed), Marcel Dekker Press, 1989.

Hlastala, M.P., and Berger, A.J., *Physiology of Respiration,* Oxford University Press, New York, N.Y., 1996.

Horsefield, K. *Anatomical Factors Influencing Gas Mixing and Distribution,* in Gas Mixing and Distribution in The Lung, (Engel, L.A., and Paiva, M., Editors), Vol 25, Lung Biology in Health and Disease, Dekker Press, 1985.

International Commission on Radiological Protection (ICRP), *Report of the Task group on Reference Man,* pp. 273-285, Oxford Pergamon, 1975.

Jacquez, J.A., *Compartmental Analysis in Biology and Medicine,* University of Michigan Press, Ann Arbor, Michigan, 2nd Edition, 1985.

Kaplan W., *Advanced Engineering Mathematics,* Academic Press, 1981.

Kreyszig, E., *Advanced Engineering Mathematics,* John Wiley, New York, N.Y., 6th ed. 1988.

Landahl, H.D., and Herrmann, R.G., *Retention of Vapors and Gases in the Human Lung and Nose,* Arch. Ind. Hyg. Occup., vol.1, pp.36-45, 1950.

Lerou, J.G.C., Dirksen, R., Kolmer, H.H.B., and Booji, L.H.D.J., *A System Model for Closed-Crcuit Inhalation Aneshesia,* Anes. , vol 75, pp.345-355, 1991

Lowe, H.J., and Ernst, E.A., *The Quantitate Practice of Anesthesia: The Use of the Closed Circuit,* Williams and Wilkens Publishers, Baltimore, MD., 1981.

Lowe, H.J, *Dose-Regulated Penthrane Anesthesia,* Abbott Laboratory, Chicago Illinois, 1972.

MacDonald, N., *Time Lags in Biollogical Models,* Springer-Verlag, 1978.

MATLAB Reference Guide, The Mathworks Inc., Natick, Mass., 1992.

Mapleson, W.W., *Circulation-time Models of the Uptake of Inhaled Anesthetics and Data for Quantifying Them,* Brit. J.Anaesth. Vol 45, pp. 319-333, 1973.

Medinsky, M.A., Kimbell, J.S., Morris. J.B., Gerde, P., and Overton, J.H., *Advances in Biologically-Based Models for Respiratory Tract Uptake of Inhaled Volatiles,* Fundam. Appl. Toxical. , vol 20, pp. 265-272, 1993.

Munson, E.S., Eger, E.I., and Bowers, D.L., *Effects of Anesthetic-depressed Ventilation and Cardiac output on Anesthetic Uptake: A Computer Nonlinear Simulation,* Anes., vol. 35, pp. 251-259, 1973

Oldendorf, W.H., *Measurement of the Mean Transit Time of Cerebal Circulation by External Detection of an Interveneously Injected Radioisotope,* J. Nucl. Med., vol. 3, pp. 382-398, 1962.

Patankar, S.V., *Numerical Heat Tramsfer and Heat Flow,* Hemisphere Publishing Co., New York, N.Y., 1980.

Pedley, T.J., Schroter,R.C., and Sudlow, M.F., *Gas Flow and mixing in the Airways,*in Bioengineering aspects of the Lung, J.B. West, ed., . Marcel Dekker Inc., Chapter 3, pp. 163-265, 1977.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling W.T. , *Numerical Recipies: The Art of Scientific Computing,* Cambridge University Press, Cambridge, U.K., 1986.

Riggs, D.R., and Goldstein, A., *Equation for Inert Gas Exchange which Treats Ventilation as Cyclic,* J. Appl. Physiol., vol. 16,pp531-537, 1961.

Scherer, P.W., Shendalman, L.H., and Greene, N.M., *Simultaneous Diffusion and Convection in Single Breath Washout,* Bull. Math. Biophysics vol 34, pp. 393-412, 1972.

Schrikker, A.C.M., de Vries, W.R., Zwart, A., and Luijendijk, S.C.M., *Uptake of Highly Soluble Gases in the Epithelium of the Conducting Airways,* Pflügers Archiv (Europeon Jourmnal of Physiology), vol. 405, pp,389-394, 1985.

Smith, N.T., Eger, E.L, Stoelting, R.K., and Whitcher, C.E., *Cardiovascular Effects of Halothane in Man: Studies During Induction and Sudden Changes in Concentration,* JAMA, vol 206, pp.1495-1499, 1968.

Smith, N.T., Zwart, A., and Beneken, J.E.W., *Intreraction Between Circulatory Effects and the Uptake and Distribution of Halothane,* Anes., vol 37, pp. 47-58,1972.

Stoelting, R.K., *Halothane and Methoxyflorane Concentrations in End-tidal Gas, Arterial Blood, and Lumbar Cerebrospinal Fluid,* Anes., vol 38, pp. 384-388, 1973.

Vinegar, A., Willaiams, R.J., Fisher, J.W., and McDougal, J.N., *Dose-Dependent Metabolism of 2,2-Dichloro-1,1,1-trifloroethane: A PhysiologicallyBased Pharmacokinetic Mosel in the Male Fischer 344 Rat,* Toxicol. Appl. Pharm., vol. 129, pp. 103-113, 1994.

Vinegar, A., and Jepson, G.W., *Cardiac Sensitization Thresholds of Halon Replacement Chemicals Predicted in Humans by Physiologically Based Pharmacokinetic Modeling,* Toxicol. Appl. Pharm., 1996, (In Press).

Wagner, P.D. *Peripheral Inert-Gas Exchange,* Chapter 14, pp. 257-281, Handbook of Physiology, Section 3, the Respiratory System, (Fishman, A.P., Farhi, L.E., and Geiger, S.R., editors)American Physiological /society, Bethesda, MD., 1987.

Watson, E.J., *Diffusion in Oscillatory Pipe Flow,* Journ. Fluid Mech., vol 133, pp233-244, 1983.

Weibel, E.R., *Morphmometry of the Human Lung,* Academic Press. PP. 110-1453, 1963.

West, J.B., *Respiratory Physiology,* Williams and Wilkens Press, Baltimore, MD., 4th Ed., 1990.

Willaiams, R.J., Vinegar, A., McDougal, J.N., Jarabek, A.M. and Fisher, J.W., *Rat to Human Extrapolation of HCFC-123 Kinetics Deduced From Halothane Kinetics: A Corollary Approach to Physiologically Based Pharmacokinetic Modeling,* Fund. Appl.Toxicol., vol 30, pp. 55-60, 1996.

Zou, Q., Liu, Z., and Goldberg, I.S. *On Nonaxisymmetric Entry Flow at Very Low Reynolds Numbers,* Mathematical Biosciences, vol. 113, pp. 245-260, 1993.

INVESTIGATION OF THE SUITABILITY OF TACTILE
AND AUDITORY STIMULI FOR USE IN
BRAIN ACTUATED CONTROL

Robert J. Hirko
Associate Engineer
Department of Aerospace Engineering,
Mechanics, and Engineering Science

231 Aerospace Building
University of Florida
Gainesville, FL   32611-6250

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

August 1966

INVESTIGATION OF THE SUITABILITY OF TACTILE AND AUDITORY STIMULI FOR USE IN BRAIN ACTUATED CONTROL

Robert J. Hirko
Associate Engineer
Department of Aerospace Engineering,
Mechanics, and Engineering Science
University of Florida

Abstract

Brain actuated control, BAC, has been demonstrated to be possible utilizing an oscillating visual stimulus, acting like a carrier signal, which is modulated by thought processes and detectable in the EEG of a trained subject. This modulated EEG signal has been used to control a variety of physical devices. Among other applications, this type of "hands off" control could be eminently applicable to the assistance of people with disabilities to let them control their environment. The visual stimulus required, however, may not be appropriate for use out of a well controlled space. For this study we investigated alternate stimuli to evaluate their possible use in a similar BAC operating mode. Tactile and auditory stimuli were used at the same fundamental frequency as the visual stimulus, 13.25 Hz, and also at a higher frequency, 43 Hz. The specific goals were to successfully detect the stimulus carrier in the EEG signal as a steady state evoked response, and to determine best locations for electrodes.

Results of the study show that stimulation by both auditory and tactile means generated evoked responses that were of useable amplitude. Fundamental frequency components as well as some second and third harmonics were observed. Power spectral levels of the fundamental were routinely 5 or 6 dB above the local baseline for the auditory and tactile stimuli. In certain cases as much as 9 dB can be demonstrated. As a comparison, the visual stimulus results in typical fundamental frequency power spectral levels of 12 to 14 dB above baseline. Of the electrode sites investigated, electrodes giving a differential between f4 and p3 were consistently the best in this study. The positive results of this study indicate that future development of these alternative stimulus modes for BAC would have a high probability of success.

INVESTIGATION OF THE SUITABILITY OF TACTILE
AND AUDITORY STIMULI FOR USE IN
BRAIN ACTUATED CONTROL

Robert J. Hirko

Introduction

The brain of a human being is a wondrous mix of substances structurally and functionally organized in such a manner as to provide electrochemical reactions that permit the human to sense and control his environment in a way that no machines have yet matched in scope. Even computers, as fast as they have become, cannot match the wide bandwidth capability of the brain-visual system pair for high speed pattern recognition and decision making. It is this superior capability of the human brain that has invited attempts to develop interfaces which would permit the direct utilization of brain function in the control of external devices. Applications for an interface of this sort exist in many areas. The Air Force envisions the possibility of reducing pilot workload with such a device. Control capability through thought processes would also have a marked effect on the lives of those people with limited physical mobility. Even in the commercial entertainment market for virtual reality devices, which is expanding rapidly, use of direct brain interface devices would play a significant part.

Significant achievements in pioneering work on brain actuated control (BAC) are being made at Armstrong Laboratory (AL/CFHP) by McMillan et. al. [1,2] utilizing modulation of electroencephalographic responses to a stimulus signal. This modulation of the "carrier" (i.e. enhancement or suppression of response magnitude) is actively performed by thought processes of the trainee, and the amount of modulation is used as a control signal. They have demonstrated success at controlling several tasks with their system, including rudimentary control of functional electrical stimulation (FES) of thigh muscles for knee extension.

Problem Discussion

While McMillan and his group are demonstrating the capability of BAC utilizing a visual stimulus in the controlled environment of the laboratory, it is questionable whether a visual

stimulus system is appropriate for use out in the real world. Interest in using this system outside of the laboratory exists both within the Air Force community and with those of us working with disabled individuals. The potential is high that a system based on BAC can be developed which would permit persons with limited function to control their environment. Stimuli other than visual, such as auditory or tactile, may be more convenient to use while moving from place to place or room to room. Therefore the purpose of this study was to investigate these alternative stimuli, and evaluate their suitability for use in a BAC system.

The BAC experimental setup at AL/CFHP utilizes visual steady state evoked potentials (SSEP) as the detected signals for the control experiments. Visual SSEPs are strong signals and relatively easy to demonstrate. It was not clear that the same result would be found looking for tactile and auditory SSEPs. Much has been written in the literature regarding auditory evoked potentials (AEP) [3,4,5,6,7,8] and somatosensory evoked potentials (SEP) [9,10,11,12,13,14,15,16,17,18], which are generated by tactile stimulation. The EPs described by these authors are, however, transient signals, and synchronous averaging of many sequential EPs is necessary to bring them out of the EEG 'background'. Utilization of these EPs incorporates analysis of any of the various latent waves which make up the time history of the transient. Very little in the literature has been written regarding SSEPs seen from somatosensory stimulation, and those articles describing auditory SSEPs are mostly related to clinical evaluation [7] and usually not within the stimulus frequency range of interest for BAC as achieved today. However a common thread that is seen in all the references on tactile stimulation indicates that steady state EPs are possible from somatosensory stimuli. In addition a well done study by Namerow et.al. [10], which showed an almost steady state response to limited length somatosensory stimulus trains, offered encouragement that SSEPs were possible with tactile stimuli.

Methodology

In searching for auditory and tactile SSEPs it seemed reasonable to assume that the potentials, if found at all, would be detectable in the same areas of the scalp where transient EPs are found. It was fortunate that Goff et. al. [9] reported on scalp potentials for human somatosensory and auditory EPs measured using the standard 10-20 International System for electrode placement. This gave us a starting point and a comparison for our findings.

Our initial evaluations to identify good candidates for electrode sites were performed using a Biologic clinical EEG unit with Brain Atlas software. Full 10-20 scalp recording was done on one test subject that had previously showed good responses to visual stimulation. The intent in using that specific subject was to have someone whose visual data was on file for comparison, and who had a high probability of being able to generate responses to tactile and auditory stimuli.

Stimulation was accomplished in two ways, both for these initial evaluations and for the later BAC trials. The auditory stimulus was a bilateral auditory rectangular pulse of one millisecond duration at the repetition rate of the fundamental frequency of stimulation. Auditory level was adjusted for approximately 80 dB. For the majority of trials in this study stimulus frequencies were 13.25 Hz and 43 Hz. Tactile stimulation was accomplished with a Brüel & Kjaer mini shaker type 4810. It was driven by a train of square pulses of 10% duty cycle at the fundamental frequency. The tactile unit pulsed the nail of the middle finger of the right hand. Intensity of this pulsing was not measured but subjectively adjusted to be comfortable for the trainee for an extended trial session.

Initial evaluations using the Biologic Brain Atlas were done to select electrode sites for use during later tests in the BAC laboratory facilities. Most of the data from these runs showed promising signals in frontal, central, and parietal areas, and were in general agreement with published potentials of Goff et.al.[9]. From these initial data we chose a limited set of electrodes to use when evaluating the stimuli in the BAC laboratory with the complete feedback and control system. There were two reasons for using the BAC system. First we could record and analyze more parameters of the signals than we could with the Brain Atlas, and second we could obtain a preliminary evaluation of the alternate stimuli in a control mode trial. EEG signals were collected with reference to right and left mastoid electrodes. The right was always the reference electrode and the left was always ground. The EEG signals which were collected and analyzed were at two other sites, each recorded with reference to the right mastoid. The system then generated a difference between these two EEG signals, and this differential signal was used as the control signal in the BAC simulation. Pairs of sites investigated were (fz,cz), (f4,f3f7), (fpz,c3t3), (p3,cz), and (f4,p3). Locations designated as (f3f7) or (c3t3) were located halfway between the standard f3 and f7 locations or halfway between c3 and t3.

Data that was recorded from each of the five pairs of sites was analyzed for frequency component energies to determine the extent to which the stimulus was present in the EEG. During the BAC trials other data was also compiled, such as commands to enhance or suppress the SSEP and the resulting effect on the device being controlled, so that we could compare, to some degree, performance of the tactile and auditory stimuli with results acquired from visual stimulation. Also fundamental, second, and third harmonic spectral energy during specific enhancement command segments was compared to that during suppress command segments to demonstrate control capability.

Results

Examples of raw spectral data for the BAC trials are shown in Figures 1 and 2. Figure 1 shows the complete spectrum, from 0 through 80 Hz, for a tactile stimulation trial with a stimulus of 13.25 Hz. It happens to be the spectrum for the complete set of data taken during the trial which showed the best learned control for this series. Data from lower scoring trials showed similar spectral behavior. Note the energy component at the fundamental frequency. It indicates that we can get signals from tactile stimulation which are as much as 6 dB above the baseline. This data shows a clear second harmonic at 26.5 Hz, a third harmonic at 39.75 Hz, and the 40 Hz filter rolloff.

Figure 2 is an example of a 43 Hz tactile stimulus trial. It clearly shows that the primary signal is present at a level of 6 dB above the baseline. Upon closer inspection in an expanded plot as shown in Figure 3, it is apparent that the signal is even greater, possibly as much as 9 dB. To make recording possible at 43 Hz the cutoff filters which normally are set at 40 Hz for the BAC system had their cutoff frequencies shifted to $11\overline{0}$ Hz. This permits a large 60 Hz component and its subharmonics to be seen as in Figure 2. Figures 1 through 3 are characteristic of all the tactile data collected. Auditory stimulion resulted in similar signals and levels.
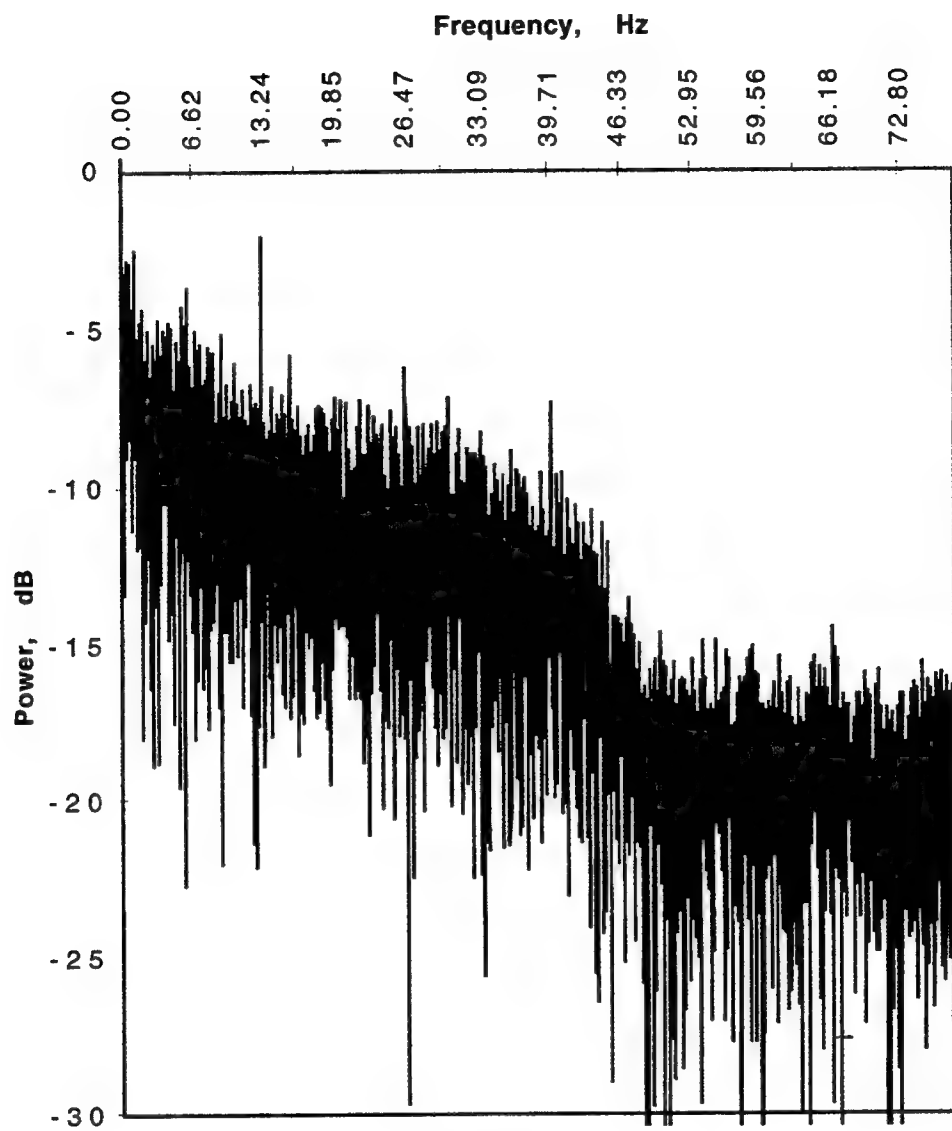
Figure 1
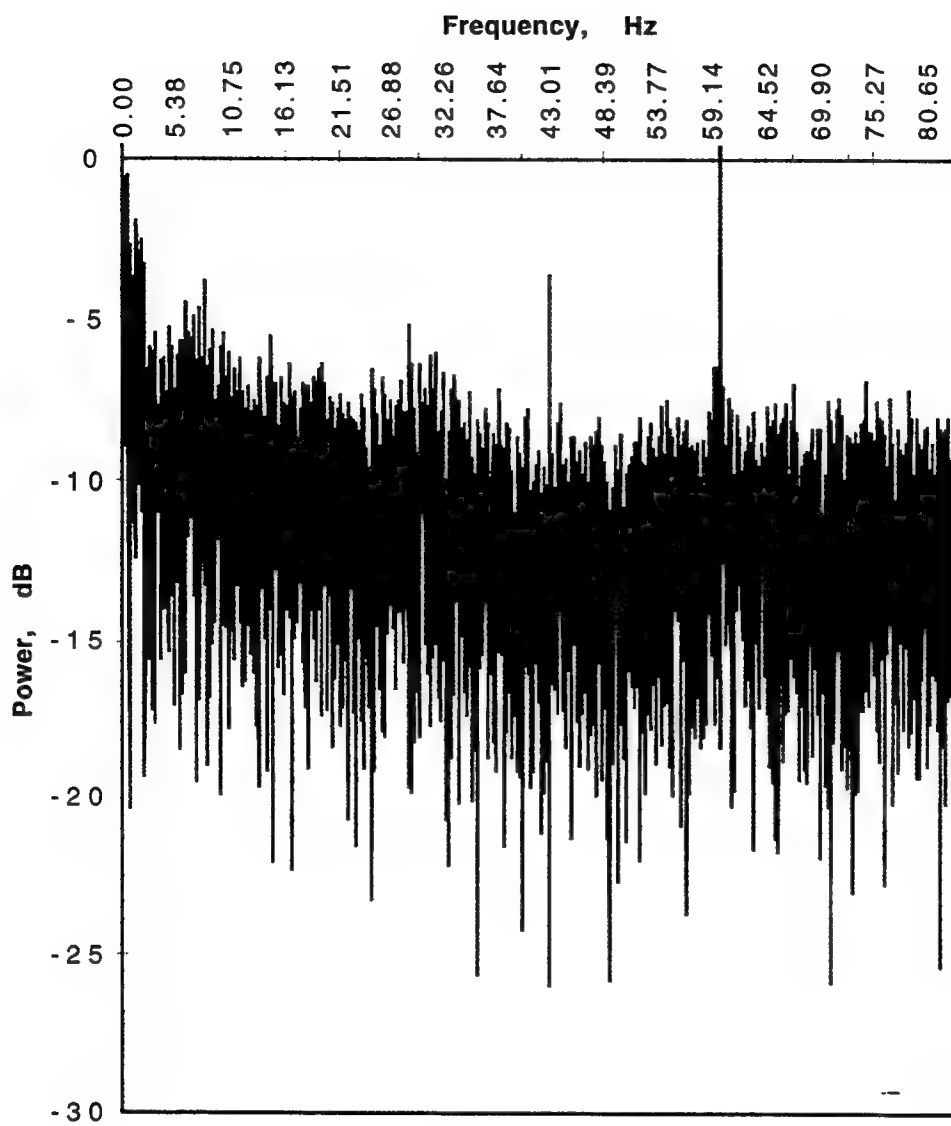
Figure 2
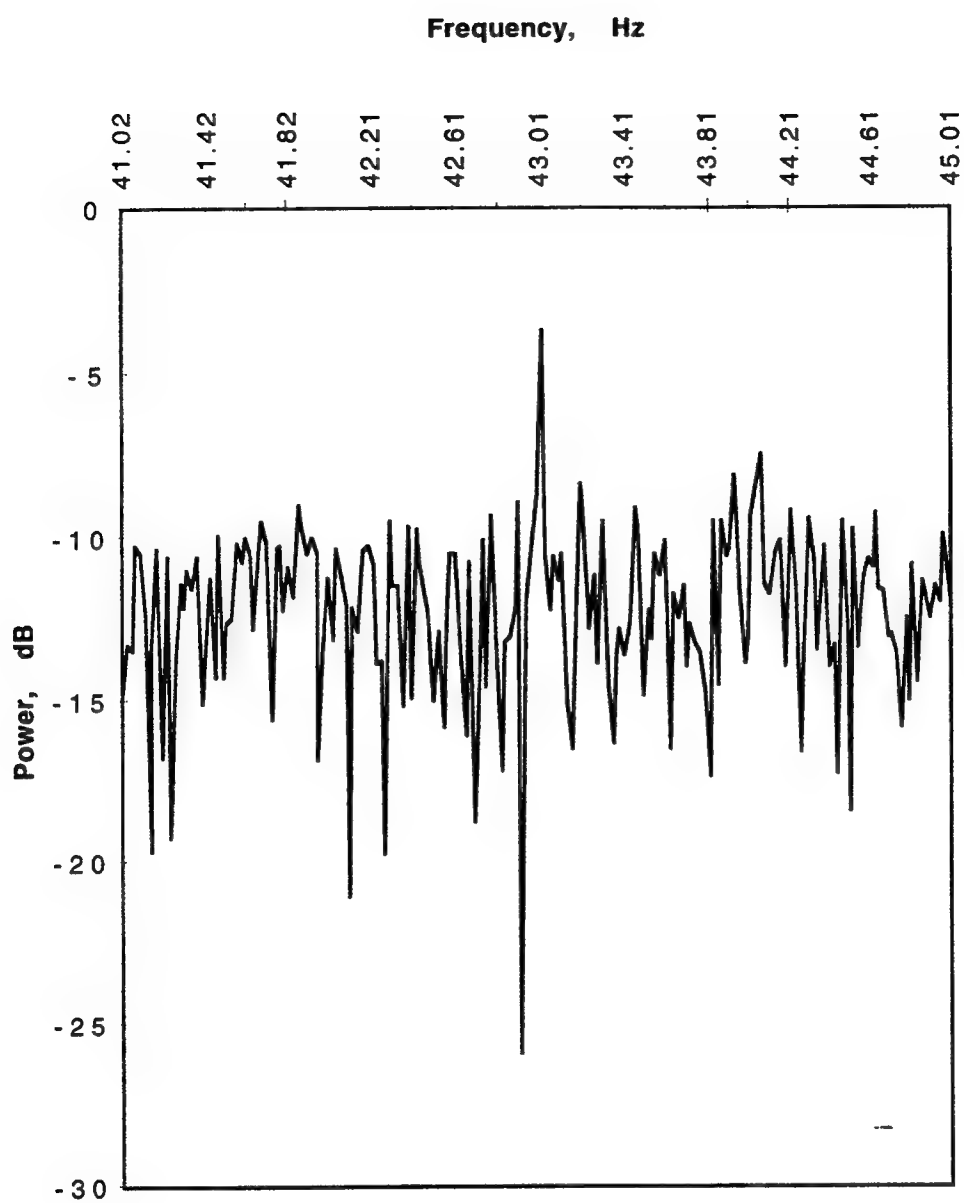
Figure 3

During the course of each BAC trial the subject may have had several enhance or suppress command tasks where the intent was to either increase or decrease the SSEP response to the stimulus. These segments of a trial were identified and the power spectrum calculated. Clear differences were seen between enhance and suppress segments. Figures 4 and 5 illustrate this effect for an auditory stimulus.

Figure 4 illustrates a 10 Hz wide portion of the power spectrum centered at the stimulus frequency for an enhance segment. A grid line is plotted at 13.28 (which is approximately the stimulus frequency of 13.25) for ease of comparison. The peak seen here is -4 dB at the stimulus frequency. Figure 5 shows a similar power spectrum plot for a suppress segment during the same auditory stimulus trial. A valley of -10 dB can be seen at the stimulus frequency. A clear 6 dB difference such as this is enough of a change to use for control purposes.

These figures illustrate the characteristics found in all the data collected with auditory and tactile stimulation during the BAC trials. As a comparison, Figure 6 is included to illustrate a typical power spectrum of raw EEG data from a trial using the same subject however with a visual stimulus at 13.25 Hz. The fundamental response is clearly stronger than that which was seen with the other stimuli in this study. A peak of approximately 12 to 14 dB above the baseline is seen.

In all, over sixty BAC trials were run during the course of this study. Variations in stimulus frequency and EEG electrode locations were made. Frequency analysis of all trials was performed and desired signal components could be found in most data that was accumulated. Since this data was collected under the control of a BAC task, the relative performance could also be evaluated in a sense. Due to the variations in experimental conditions we did not get the numbers of trials, under any one condition, needed for statistical evaluations, but an interesting observation was made nevertheless. When the data was evaluated for average signal level above enhance threshold during an enhance segment of a trial and average signal level below suppress threshold during respective suppress segments, it was seen that the difference between enhance average and suppress average tended to increase as the study went on. This would seem to indicate that learning was taking place from one trial to the next.
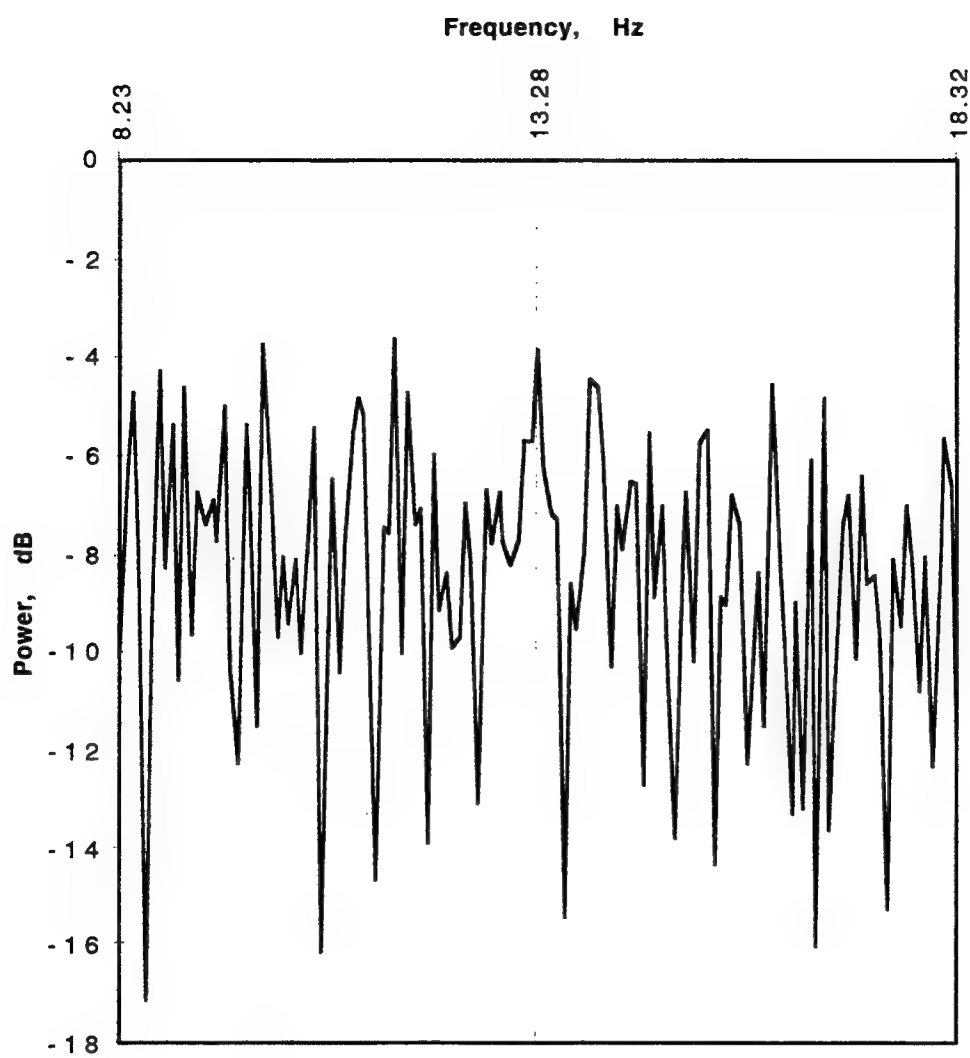
Figure 4
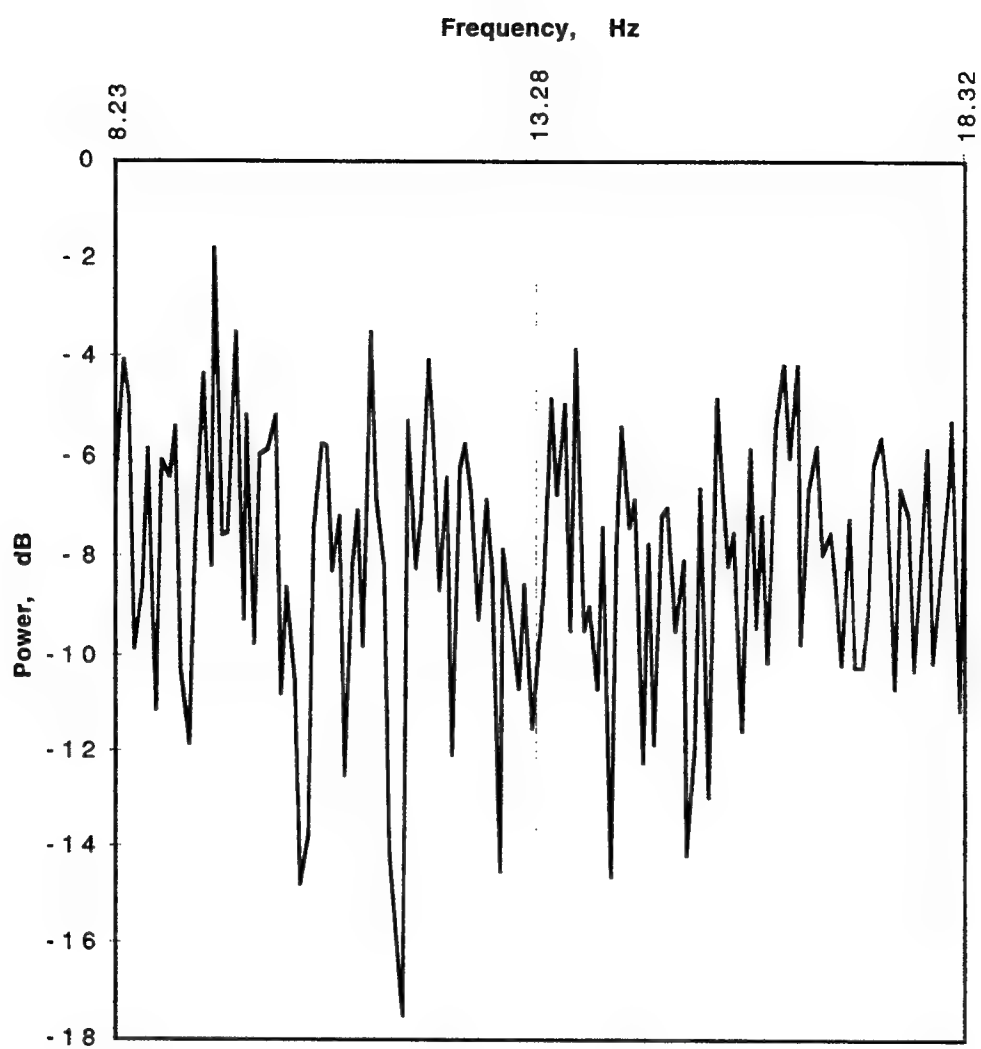
Figure 5
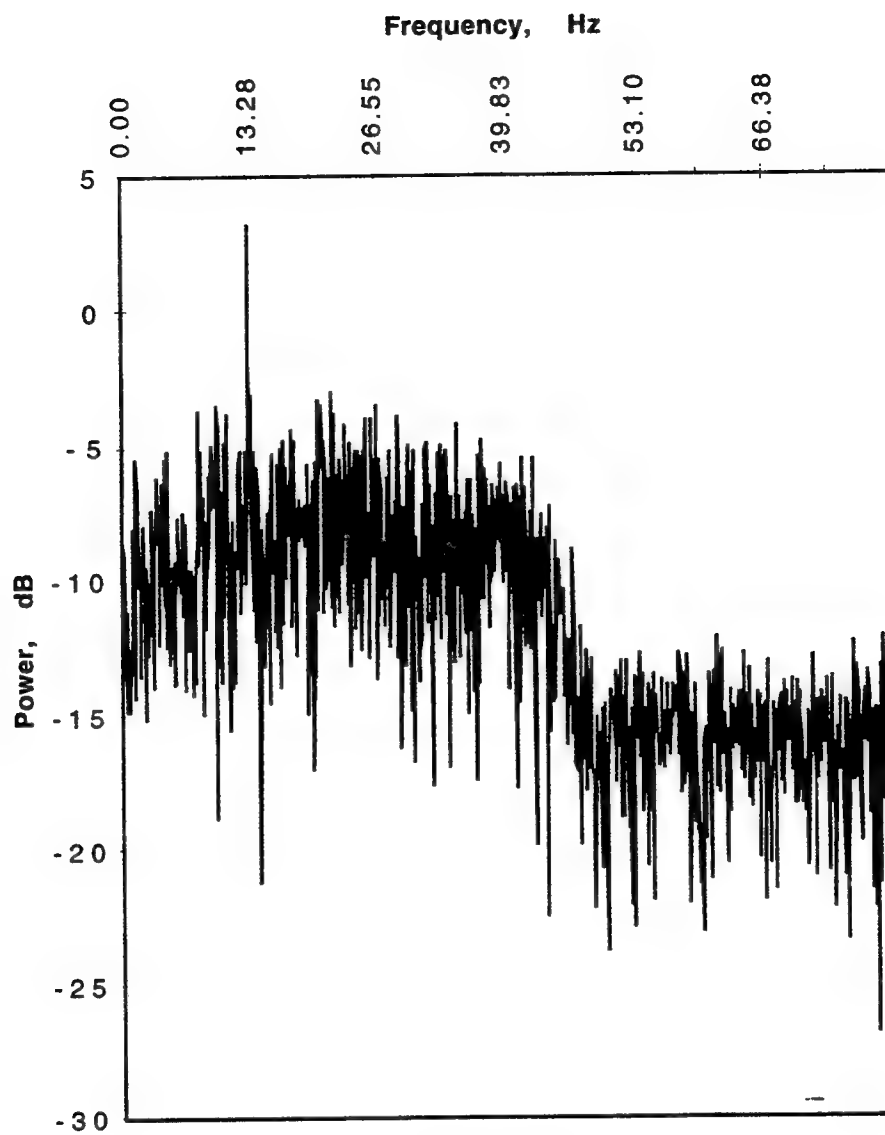
Figure   6

Conclusion

This study was successful at meeting its primary goals of finding evoked responses and identifying suitable EEG electrode sites for auditory or tactile BAC trials. We routinely demonstrated 5 to 6 dB of fundamental frequency signal available for use in the BAC studies. We even demonstrated as much as 9 dB, significantly approaching that level seen with visual stimulus. This signal was seen to more or less of a degree over all the electrode sites investigated, but we had best success with the (f4,p3) pair. Even though this site was successful for us, it is not clear that the observed success in BAC learning is the best we can achieve from optimum site selection and experimental trial design. Nor is it clear that the successful control is a result of the same mode of brain control that is observed in the visual stimulus studies. There, enhancement results in a shift of SSEP magnitude between hemispheres, the degree of which is utilized as the control signal. The visual-stimulus EEG sites are (o1,o2), spanning the midline. Since we were looking across the midline and across the central sulcus with our sites for detecting the auditory or tactile signals, the type of signal shift is clearly not the same. Thus the question of just what was causing the differential signals that we detected is one that needs to be answered in future investigations. Understanding that will help us choose an optimal detection scheme and possibly develop detected signals at the 12 to 14 dB levels that we see with the visual stimulus.

Another interesting observation that needs to be fully explained and utilized, if a real phenomenon, is the 2nd and 3rd harmonics which were frequently seen in both tactile and auditory trials. Indeed the third harmonic was commonly of the same magnitude as the fundamental and occasionally larger. A limited investigation of the BAC instrumentation was performed to determine whether it was generating the harmonics and no evidence to suggest that was found. If these are real signal components the question as to whether they can offer increased control capability needs to be asked.

In the end it is clear that these two alternate stimuli for BAC are viable and warrant more investigation. Future work must be done in the signal and instrumentation area as well in the design of psycophysiological testing for evaluation and control design. Proposals for this work will be forthcoming.

References

1. McMillan, G.R., Calhoun, G.L., Middendorf, M.S., Schnurer, J.H., Ingle, D.F., & Nasman, V.T. (1995). Direct brain interface utilizing self-regulation of the steady-state visual evoked response. *Proceedings of the RESNA 18th Annual Conference*, 693-695.

2. Calhoun, G.L., McMillan, G.R., Morton, P.E., Middendorf, M.S., Schnurer, J.H., Ingle, D.F., Glaser, R.M., & Figoni, S.F., (1995). Functional electrical stimulator control with a direct brain interface. *Proceedings of the RESNA 18th Annual Conference*, 696-698.

3. Azzena, G.B., Conti,G., Santarelli,R., Ottaviani,F., Paludetti,G., Maurizi,M. (1995). Generation of human auditory steady-state responses(SSRs).I: Stimulus rate effects. *Hearing Research* 83, 1-8.

4. Santarelli,R., Maurizi,M., Conti,G., Ottaviani,F., Paludetti,G., Pettorossi,V.E., (1995). Generation of human auditory steady-state responses(SSRs).II: Addition of responses to individual stimuli. *Hearing Research* 83, 9-18.

5. Makeig,S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology*, 86, 283-293.

6. Pantev, C., Elbert, T., Makeig, S., Hampson, S., Evlitz, C., Hoke, M. (1993). Relationship of transient and steady-state auditory evoked fields. *Electroencephalography and Clinical Neurophysiology*, 88, 389-396.

7. Rickards, F.W., Tan, L.E. Cohen, L.T., Wilson, O.J., Drew, J.H., Clark, G.M. (1994). Auditory steady-state evoked potential in newborns. *British Journal of Audiology*, 28, 327-337.

8. Dobie, R.A., Wilson, M.J. (1994). Phase weighting: A method to improve objective detection of steady-state evoked potentials. *Hearing Research* 79, 94-98.

9. Goff, G.D., Matsumiya, Y., Allison, T., Goff, W.R. (1977). The scalp topography of human somatosensory and auditory evoked potentials. *Electroencephalography and Clinical Neurophysiology*, 42, 57-76.

10. Namerow, N.S., Sclabassi, R.J., Enns, N.F. (1974). Somatosensory responses to stimulus trains: normative data. *Electroencephalography and Clinical Neurophysiology*, 37, 11-21.

11. Larsson, L.E., Prevec, T.S. (1969). Somato-sensory responses to mechanical stimulation as recorded in the human EEG. *Electroencephalography and Clinical Neurophysiology*, 28, 162-172.

12. Ishiko, N., Hanamori, T., Murayama, N. (1980). Spatial distribution of somatosensory responses evoked by tapping the tongue and finger in man. *Electroencephalography and Clinical Neurophysiology*, 50, 1-10.

13. Kunde, V., Treede, R.D. (1993) Topography of middle-latency somatosensory evoked potentials following painful laser stimuli ans non-painful electrical stimuli. *Electroencephalography and Clinical Neurophysiology,* 88, 280-289.

14. Nakanishi, T., Takita, K., Toyokura, Y. (1973). Somatosensory evoked responses to tactile tap in man. *Electroencephalography and Clinical Neurophysiology,* 34, 1-6.

15. Pratt, H., Amlie, R.N., Starr, A. (1979). Short latency mechanically evoked somatosensory potentials in humans. *Electroencephalography and Clinical Neurophysiology,* 47, 524-531.

16. Starr, A., Mckeon, B., Skuse, N., Burke, D. (1981). Cerebral potentials evoked by muscle stretch in man. *Brain,* 104, 149-166.

17. Purves, A.M., Boyd, S.G. (1993). Time shifted averaging for laser evoked potentials. *Electroencephalography and Clinical Neurophysiology,* 88, 118-122.

18. Beydoun, A., Morrow, T.J., Shen, J.F., Casey, K.L. (1993). Variability of laser-evoked potentials: attention, arousal, and lateralized differences. *Electroencephalography and Clinical Neurophysiology,* 887, 173-181.

DETERMINATION OF THE INFLUENCE OF ULTRAWIDEBAND EXPOSURE OF RATS DURING EARLY PREGNANCY ON PREGNANCY RATE, EMBRYONIC SURVIVAL TO TERM, AND SEX RATIO OF OFFSPRING

Michael P. Dooley, MS, PhD
Affiliate Assistant Professor of Physiology
Department of Veterinary Physiology and Pharmacology
College of Veterinary Medicine

Iowa State University
2032 Veterinary Medicine
Ames, IA 50011-1250

DETERMINATION OF THE INFLUENCE OF ULTRAWIDEBAND EXPOSURE OF RATS DURING EARLY PREGNANCY ON PREGNANCY RATE, EMBRYONIC SURVIVAL TO TERM, AND SEX RATIO OF OFFSPRING

Michael P. Dooley, MS, PhD
Affiliate Assistant Professor of Physiology
Department of Veterinary Physiology and Pharmacology
College of Veterinary Medicine
Iowa State University

Abstract

Several epidemiological surveys have detected statistically significant associations between health effects and residential and/or occupational exposures to electromagnetic fields. Field effects have been reported for exposures ranging from fields which approach DC fields, including extremely-low-frequency fields (<300 Hz), particularly those at 50 and 60 Hz, to high frequency radiation in the kHz to MHz ranges. A mechanism or mechanisms that would explain such effects have not been developed although some hypotheses have been formulated that could provide a biological basis to explain the observed outcomes. Among these reports, Larsen et al. (1991) found that occupational exposure of Danish physiotherapists to non-ionizing radiation was associated with alterations in gender ratio and birth weight of offspring. In an attempt to evaluate experimentally whether such an effect could be demonstrated, studies were initiated using a rat model. The UWB source used for this study was a microwave transmitter that was capable of generating electromagnetic pulses ranging from near DC to several GHz, but had, as a primary component, microwave pulses in the MHz range. Virgin female rats were paired with males of known fertility. Mated females were then randomly assigned to a sham-treatment or to an UWB-exposure group. Rats were exposed to the assigned treatment during the first 3 days of pregnancy. Females were then monitored for pregnancy, birth and survival of offspring, and sex distribution of the pups born. No effects of electromagnetic field exposure were seen for pregnancy rate, postimplantation embryonic survival, birth of live offspring, or gender ratio.

DETERMINATION OF THE INFLUENCE OF ULTRAWIDEBAND EXPOSURE OF RATS DURING EARLY PREGNANCY ON PREGNANCY RATE, EMBRYONIC SURVIVAL TO TERM, AND SEX RATIO OF OFFSPRING

Michael P. Dooley, MS, PhD

Introduction

Exposure to various forms of electromagnetic radiation is inherent to life within industrialized societies. In recent years, increasing numbers and types of radiowave and microwave transmitters have been developed and others are being developed with the potential for use in military and/or civilian applications. These systems produce relatively unique forms of electromagnetic field exposure, including electromagnetic pulses at high frequencies, with rapid risetime, and in the case of UltraWideBand (UWB) systems, result in the simultaneous emission of electromagnetic pulses over a broad frequency spectrum.

Although microwave radiation effects on biological systems have been shown for exposure systems in which heating of cells and tissues can be documented, there has not been consensus or cell and animal models developed with the capacity to discriminate between thermal and athermal effects in biological systems. Such models are needed to provide an experimental basis to confirm or refute the numerous associations which appear to exist between certain forms of electromagnetic radiation and biological responses and/or human health consequences.

Maintenance of reproductive functions is essential to the continued survival of the species. Thus, factors affecting reproductive processes which have the potential to exert profound effects in either artificial or natural biological systems need to be identified.

As a result of studies performed in our laboratories at Iowa State University, we have proposed (Dooley, **et al.,** 1994; Lamont **et al.,** 1994) that mammalian embryos at the preimplantation stages offers an animal model system in which the cellular effects of exposure to electromagnetic fields can be determined. In this system, exposures are applied either to embryonic cells during defined periods of **in vitro** culture, while the embryos are developing **in vivo** within the mother, or **in vivo** within intermediate hosts. Because the continued viability and development of these embryonic cells are integral to the survival of the species, they may provide a sensitive and biologically relevant indicator to assess for detrimental consequences of electromagnetic field exposure, particularly for those forms of electromagnetic radiation that are generated as a result of human activity.

In recent years several epidemiological surveys have detected statistically significant associations between biological effects and residential and/or occupational exposures to electromagnetic fields ranging from extremely-low-frequency fields (<300 Hz), particularly those at 50 and 60 Hz, to high frequency radiation in the kHz to MHz range. Bioeffects which have been linked to exposure to electromagnetic radiation include an increased incidence of cancer in children, reduced fertility in humans, alteration of behavioral responses in rats and humans, developmental abnormalities in chickens, mice, and rats, and a reduced rate of development in fish and mammalian embryos (Frey, 1993; Zusman **et al.,** 1990). Among the various surveys was a report (Larsen **et al.,** 1991) that occupational exposure of Danish physiotherapists to non-ionizing radiation was associated with alterations in gender ratio and decreased birth weight of male offspring. Larsen **et al.** (1991) reported that the association was stronger for female physiotherapists in which the exposure estimates were the highest. To our knowledge experimental studies to test this hypothesis have not been performed although a survey performed by Gubéran **et al.** (1994) did not find any association for such an effect among Swedish physiotherapists. Recently, James (1995) outlined a possible

biological basis which could explain such a result, if in fact, it could be replicated.

This study was performed to determine whether daily exposures of female rats to an UWB electromagnetic field on post-mating days 0, 1, and 2 affected the viability and **in vivo** development of embryos and/or the sex distribution of offspring born to exposed females. Specific objectives were to: a) confirm that preimplantation stage embryos exposed to UWB pulses were capable of implanting and developing to term and b) determine the sex-ratio of the offsprings born following UWB exposure of pregnant females.

Materials and Methods

Animals

Sprague-Dawley-derived male and female rats were used for this study. A total of 8 male rats (> 300 g, > 100 days of age) of proven fertility and 36 virgin, Sprague-Dawley-derived female rats (> 200 g, > 100 days of age) were available for use in this study. Males were caged individually and provided food and water **ad libitum.** Female rats were caged in pairs and provided food and water as for the males. All rats were housed in plastic cages, kept in the same room, and maintained at 23 °C (range 21 - 27 °C during the study) and in a 14-hour light / 10-hour dark illumination cycle.

Mating and Assignment to Treatment

Female rats were selected for mating based on vaginal cytology indicative of proestrus or estrus. Female rats were then paired with a sexually mature male rat and left in the male's cage during the dark phase of the light cycle (14 hours' light; 10 hours' dark). Mating was confirmed the following morning (day 0) by the presence of spermatozoa in the vaginal smear (Dooley, 1988). Females in which spermatozoa or spermatozoal remnants could be identified within at least 10 microscopic fields of the vaginal smear were considered to be pregnant. Females

judged to have mated were then randomly assigned to a treated (UWB-exposed) or control (sham-exposed) group. However, for each pair of rats, the remaining female in that cage was then allocated to the other treatment group (UWB-exposed or sham-exposed) to ensure that comparable numbers of rats would be assigned to each group.

UWB Exposure

A Kentech microwave transmitter was used to generate microwave pulses at multiple frequencies in the MHz range. The pulse characteristics were as follows:

Peak E-field intensity = 61 kV/m
Pulse risetime = 300 ps
Pulse width = 1.8 ns
Pulse rate = 1 kHz.

The output of the Kentech source was controlled using a commercial pulse generator (Hewlett-Packard 8112A Pulse Generator 50 MHz). The settings used with the Hewlett-Packard pulse controller for these studies were as follows:

MODE = Norm
CTRL = Off
PER = 1.00 ms (1000 Hz)
DEL = 65.0 ns
WID = 100 μs
DBL = 200 μs
DTY = 50%.   --

During each and every animal exposure, the output of the Kentech source was continuously monitored by the detection of an E-field signal which was generated between the parallel transmission plates of the exposure facility whenever the pulsed output of the Kentech source was activated. This signal was monitored using an oscilloscope, such that the oscilloscope amplified and displayed an electronic signal that was representative of the E-field generated between the parallel plates of the transmitter. This oscilloscope display was therefore used to verify that the Kentech system was functional, to observe the relative shape of the

16--6

output as recorded from the parallel plates, and, with the aid of a Polaroid camera, to provide for photographic documentation of the signal generated during the daily exposures provided to UWB- or sham-treated rats to ensure that exposure conditions remained constant over the period of study.

During those periods in which the daily animal exposures were being made, the Kentech power supply, the Hewlett-Packard pulse controller, and the oscilloscope were left on at all times. The disable button of the Hewlett-Packard pulse controller was used to regulate whether pulses were applied, as for the UWB-exposed rats, or to disable the output of the Kentech system for exposures provided to the sham-treated rats. Therefore, with the exception of the UWB pulses, the exposure conditions used were identical for rats exposed to electromagnetic fields produced by these UWB pulses or during the periods of restraint provided to rats assigned to the sham-treatment group.

Each UWB and sham animal exposure was performed within an anechoic chamber. Over the course of this study, chamber temperature was regulated at 24 ± 0.2 °C during each UWB and/or sham exposure. Independent measures of surface or core body temperatures of rats were not obtained. To expose rats, animals were transferred from the animal housing facilities located in an adjacent building to the UWB exposure facility. Rats were transferred in an enclosed cage, equipped with a lid containing a filter to minimize exposure to potential airborne contaminants. Immediately prior to UWB exposure or sham treatment, the rat was placed in a Plexiglas restraint, the rear restraint plate was taped down, and then the animal was transferred to the anechoic chamber. UWB-exposed and sham-treated rats were lightly restrained within this Plexiglas holder for periods in excess of 6, but not for more than 10 minutes. Rats were positioned within the microwave exposure chamber such that the cranial portion of each rat was closest to the microwave source. The Plexiglas restraint was then positioned on a Styrofoam

support and taped to the support such that the rat would be positioned between the parallel plates of the transmitter and within the central portion of the E-field that was to be generated by the UltraWideBand exposure system.

Once each rat was positioned within the transmission facility, the door to the anechoic chamber was closed and each animal was provided a 6-minute exposure to the UWB field or to a sham procedure in which the experimental conditions were identical except no UWB pulses were applied during the restraint period. At the end of each exposure, the rat was removed from the restrainer, replaced within the transport cage, and then returned to the facility housing the rat colony. During this study, rats were exposed daily for a 6-minute period to either the UWB or sham-treatment for 3 days, beginning on post-mating day 0. Exposures were provided to female rats, such that preimplantation stage embryos present within the oviducts would be exposed during their first 3 cellular cycles (1-cell, 2-cell, and 4-cell stages). Rats were then caged individually and the pregnancy was allowed to go to term. Rats were monitored at least once daily during the period of the study. Rats for which 1 or more offspring were born at term (post-mating days 21 - 22) were classified as pregnant. Rats which had not produced any offspring by post-mating day 26 were classified as non-pregnant. Pups found dead at the time of parturition were evaluated to determine whether they had died prenatally or at the time of birth and to confirm the gonadal sex of the offspring. In this study, all pups were euthanized prior to weaning age for necropsy to determine the gonadal sex.

Initially, the surviving pups were sexed by direct observation using anatomic markers, including anogenital distance, thought to be consistent with the secondary sexual characteristics of neonatal male and female pups. Pups were then euthanized to confirm the gonadal sex of each offspring, to confirm the relative normality of the internal genitalia, and to ensure that paired reproductive structures corresponding to the phenotypic sex of the individual were present the in UWB-

exposed and sham-treated offspring.

Experimental Endpoints

The following endpoints were recorded for each female rat:

Date of mating

Body weight in grams on Day 0--Pre-exposure

Body weight in grams on Day 2--Post-exposure

Date of parturition

Number of offspring born at term

Viability of offspring born at term

Sex of offspring

Body weight, presence and number of implantation sites evident within the uterine horns of mated females.

Statistical Analyses

Chi-square analysis of ratios was used to compare the proportion of UWB-exposed and sham-treated rats remaining pregnant, the number of pups born alive for control and treated rats, and the proportion of male and female offspring born to these rats. Chi-square analysis of ratios was also used to compare the proportion of rats producing litters in which the number of female pups exceeded the number of male pups and to assess the proportion of implanted embryos that had died during the prenatal period. Analyses of variance (ANOVA) were used to compare the litter size and gestation interval for sham- and UWB-exposed rats. Statistical significance was preestablished at $P \leq 0.05$ for all planned comparisons.

Results

Of the 36 female rats available for use in this study, 33 rats were mated and assigned to treatment. From these, 17 rats were assigned to the sham-treatment group and 16 rats were assigned to the UWB-exposure group. Of the remaining 3 rats, 2 were paired with males but were not assigned to treatment due to insufficient numbers of spermatozoa within the post-pairing vaginal smear. The remaining female was also mated and had been assigned to treatment (UWB-exposure), however this rat was excluded from the study due to a malfunction of the Kentech system. Each of these 3 rats was pregnant and delivered live offspring at term.

The proportion of sham-exposed rats remaining pregnant, 16/17 (94%) was not different ($P > 0.1$; Table 1) from that of the UWB-exposure group, 15/16 (94%). The gestation intervals were comparable for both sham-exposed rats (mean $\pm$ SD = 21.5 $\pm$ 0.5 days) and UWB-exposed rats (mean $\pm$ SD = 21.4 $\pm$ 0.5 days). Gestation length ranged from 21 to 22 days for all rats. The number of offspring born ranged from 0 to 17 for the sham-exposed rats (mean $\pm$ SD = 13.1 $\pm$ 2.1) and from 0 to 17 for the UWB-exposed rats (mean $\pm$ SD = 11.8 $\pm$ 4.2). As evidenced by the magnitude of the standard deviation, there was greater variation in litter size among rats in the UWB-exposure group, however this difference was not significant ($P > 0.1$; Table 2). For rats which produced offspring at term, all of the sham-treated rats produced litters in excess of 10 pups (mean $\pm$ SD = 13.9 $\pm$ 2.1; range 11-17 pups), whereas for UWB-treated rats, 3/15 rats (20%) produced litters of 10 or fewer pups (mean $\pm$ SD = 12.5 $\pm$ 2.9; range 7-17 pups).

Table 1. Proportion of rats exposed to an UWB electromagnetic field to maintain pregnancy to term and produce live offspring.

Treatment group	Pregnant		Non-Pregnant	
	No. Rats	(%)	No. Rats	(%)
Sham-exposed	16	(94)	1	(6)
UWB-exposed	15	(94)	1	(6)

The proportion of sham-exposed rats to remain pregnant and produce offspring at term was not different (P > 0.1) than for the UWB-exposed rats.

A total of 222 pups were born to the sham-treated rats (n = 16) which remained pregnant and 220 of these were born alive (99.1 %). This compares favorably (P > 0.1, Table 3) to the 188 pups born to the UWB-treated rats which remained pregnant (n = 15), from which 187 of the pups born were alive (99.5 %). Among the offspring born, 3 of the pups born to sham-treated rats and 3 of the pups born to UWB-treated rats died during the post-natal period. Neither the pups or their remains were recovered for necropsy.

Table 2. Mean litter size at term for sham-exposed rats and rats exposed to an UWB electromagnetic field on post-mating days 0, 1, and 2.

Treatment group	Litter size		
	Mean	(±SD)	Range
Sham-exposed	13.1	(2.1)	0 - 17
UWB-exposed	11.8	(4.2)	0 - 17

Data are mean (±SD) for sham-exposed (n = 17) and UWB-exposed rats (n = 16).

The mean litter size of sham-exposed rats was not different (P > 0.1) than for the UWB-exposed rats.

Table 3. Number of pups born and survival of pups to term for sham-exposed and UWB-exposed rats during post-mating days 0, 1, and 2.

Treatment group	Alive		Dead	
	No.	Pups (%)	No. Pups	(%)
Sham-exposed	220	(99.1)	2	(0.9)
UWB-exposed	187	(99.5)	1	(0.5)

Data shown are for n = 16 litters for sham-exposed and n = 15 litters for UWB-exposed rats.

The proportion of pups born alive to sham-exposed rats was not different (P > 0.1) than for the UWB-exposed rats.

Table 4. Number of male and female pups born at term to sham-exposed and to rats exposed an UWB electromagnetic field on post-mating days 0, 1, and 2.

Treatment group	Male		Female	
	No. Pups	(%)	No. Pups	(%)
Sham-exposed	103	(47)	116	(53)
UWB-exposed	89	(49)	94	(51)

Data shown are for n = 16 litters for sham-exposed and n = 15 litter for UWB-exposed rats.

The proportion of pups born as male to sham-exposed rats was not different (P > 0.1) than for the UWB-exposed rats.

Of the 222 pups born to the sham-treated rats, 219 of these were necropsied and the gonadal sex of the pups was positively determined. For these sham-treated rats, 103 of the pups born were male and 116 were female providing a gender ratio (male/female X 100) of 89. Reproductive structures appeared normal for all of the pups born to sham-treated rats. For the UWB-exposed rats, 183 of the 188 pups born were necropsied and the gonadal sex for these was positively determined. Of these, 89 of the pups born were male and 94 were female (gender ratio = 95). As for the sham-treated rats, reproductive structures appeared normal for all pups born to these UWB-exposed mothers and no evidence was obtained to suggest hermaphrodism or an inter-sex condition had occurred in any of the pups born to the rats used in this study.

The proportion of male pups born to the sham-exposed rats (103/219; 47%) was not different (P > 0.1; Table 4) from the proportion for UWB-exposed rats (89/183; 49%).

The observed ratios for male and female pups born to the sham-exposed and UWB-exposed groups were not different from an expected ratio of 1:1 (P > 0.1).

Among sham-treated females, 8 of 16 pregnant females (50%) produced litters in which the number of males born was equal to or greater than the number of females born. For the UWB-exposed rats, 6 of 15 litters (40%) resulted in an equal or greater number of male pups as compared to female siblings, however, these differences were not significant (P > 0.1, Table 5).

Implantation sites were not evident in either uterine horn for sham-exposed (n = 1) and UWB-exposed (n = 1) rats that did not remain pregnant. Hence, for these rats, the failure to establish and maintain pregnancy to term was attributed to fertilization failure, preimplantation embryonic mortality, or to both factors occurring in the same animal.

For rats in which pregnancy was established and maintained to term, postimplantation losses were evident in 10/16 females in the sham-treated group (63%) and in 10/15 females in the UWB-treated group (67%). These responses were not different (P > 0.1). Overall, 188 of 209 implanted embryos (90%) survived to term and were born to the UWB-exposed rats and 222 of 245 implanted embryos (91%) survived to term for the sham-treated rats (P > 0.1; Table 6).

In this study, there was no evidence to suggest that either during or after the application of UWB electromagnetic fields that the UWB-exposed rats were otherwise affected by the electromagnetic fields generated or to secondary effects associated with the generation of these fields.

16--14

Table 5. Proportion of sham-treated rats and females exposed to an UWB electromagnetic field to produce litters in which the number of males born was equal to or exceeded the number of female offspring for that litter.

Treatment group	Male ≥ Female		Female > Male	
	No. Rats	(%)	No. Rats	(%)
Sham-exposed	8	(50)	8	(50)
UWB-exposed	6	(40)	9	(60)

The proportion of sham-exposed rats remaining pregnant and which produced litters in which the number of female offspring exceeded the number of male offsprings at term was not different (P > 0.1) than for the UWB-exposed rats.

Table 6. Proportion of embryos to implant and survive to term and estimates of prenatal death due to postimplantation losses for sham-exposed and UWB-exposed rats.

Treatment group	Surviving embryos		Embryonic/fetal death	
	No.	(%)	No.	(%)
Sham-exposed	222	(91)	23	(9)
UWB-exposed	188	(90)	21	(10)

Data shown are based on n = 16 litters and 245 implantations sites for sham-exposed and n = 15 litters and 209 total implantation sites for UWB-exposed rats.

The proportion of implanted embryos that produced pups at term for the sham-exposed rats was not different (P > 0.1) than for embryos of the UWB-exposed rats.

Summary and Conclusions

An association between the exposure of humans to microwave fields and the sex distribution of offspring was reported (Larsen **et al.**, 1991). Although an experimental hypothesis to explain these potential effects, if they could be replicated in humans, has been proposed (James, 1995), it has not been tested. It is most likely that if mechanisms were active to alter gender ratios, they would be exerted on the female or the gametes and/or embryos within the female. Possible explanations could include an influence on the selection of X-bearing and Y-bearing spermatozoa within the female tract or by some unknown mechanism which exerts a post-fertilization effect on the survival of male embryos.

For mated rats which were assigned to treatment in this study, fertilization failure and/or preimplantation embryonic mortality resulting in the failure to maintain pregnancy to term was evident in only 1/17 of the sham-exposed and 1/16 of the UWB-exposed rats. Among the sham-exposed rats, evidence for postimplantation losses or perinatal losses occurred in 10/16 rats (63%) and for the UWB-exposed rats, losses were evident in 10/15 rats (67%).

The results obtained for these UWB-exposed rats confirm that embryos exposed **in vivo** to these fields are capable of maintaining pregnancy to term. This result is consistent with a prior study (Dooley **et al.**, 1996). In that study, this electromagnetic field exposure system was used to apply exposures to rats on post-mating days 0, 1, and 2 and we then determined that the embryos recovered on post-mating day 3 were viable (Dooley, 1988; Dooley **et. al.**, 1989) for both UWB-exposed and sham-treated rats. Furthermore, for mated rats in which exposures were initiated on post-mating day 3 and then continued on a daily basis throughout the gestation period, no detrimental effects of UWB exposure were observed (Cobb, **et al.**, 1996).

This UltraWideBand exposure system, at the power settings used for this study, is not expected to induce heating of cells or tissues. Furthermore, in preliminary studies which have been performed with this system, no measurable thermal increases were detected within test animals. Hence, it is presumed that the cellular effects of UWB-exposure to embryonic cells, if any, would also be athermal.

Although these results were obtained at the maximum output for this exposure system, they were performed at a relatively low power intensity as compared to other electromagnetic field sources. Despite this limitation, no evidence was obtained to suggest that UWB-exposure of mated rats on post-mating days 0, 1, and 2, affects pregnancy rate or the number of viable embryos to survive to term in mated rats. Furthermore, no support is provided to suggest that exposures to electromagnetic fields during the preimplantation period of embryonic development could alter the apparent sex distribution based on phenotypic markers or affect the gonadal development of offspring which were exposed, as embryos, during the preimplantation period of development.

Thus, the results of this study do not support the hypothesis that electromagnetic pulses could affect the survival or sexual differentiation of rat embryos as has been implicated for the human (Larsen et al., 1991) when exposures to ultrasound and short-wave diathermy equipment were considered. Because the exposures provided in this study were only performed on the embryos resulting in the F_1 generation for these rats, potential effects on spermatozoal transport, survival, and/or selection prior to fertilization in mated rats, or to chronic and/or cumulative effects exerted on the oocytes which are destined to be released in subsequent estrous cycles, and other aspects of the reproductive physiology of the male and female, remain to be determined.

References

Cobb, B. L., Mason, P. A., Miller, S. A., Kosub, K. R., Jauchem, J., and Murphy, M. R., 1996. A teratologic study of UltraWideBand electromagnetic field exposure. Presented at: Eighteenth Annual Meeting of the Bioelectromagnetics Society, Victoria, British Columbia, Canada, June 9-14, 1996. Poster Session B: Behavior, Abstract P-160B. BEMS Abstract Book, p. 259.

Dooley, M. P., 1988. The use of eosin B to assess the viability and developmental potential of rat embryos. Dissertation, Iowa State University, Ames, Iowa, 1988.

Dooley, M. P., Pineda, M. H., and Martin, P. A., 1989. Single or intermittent exposure of rat embryos to eosin B does not affect in vitro development, survival after transfer, birth of live offspring, and development to weaning. 22nd Ann. Meeting Soc. for the Study of Reprod., Univ. of Missouri, Columbia, MO., August 6-9, 1989, Abstract No. 337. Biol. Reprod. Vol. 40 (Suppl. 1):160.

Dooley, M. P., Pineda, M. H., Lamont, J. W., Weber, R. J., and Moye, D. J., 1994. Continuous Exposure of rat embryos to a 1.5 g electromagnetic field (EMF) does not affect in vitro development and viability. Presented at: Experimental Biology 94, Anaheim, CA., April 24-28, 1994, The FASEB Journal, Vol. 8, No. 4, Part 1, p. A398, March 1994.

Dooley, M. P., Cobb, B. L., Merritt, J. H., and Jauchem, J. R., 1996. Development of rat embryos exposed to an UltraWideBand (UWB) electromagnetic field. Presented at: Eighteenth Annual Meeting of the Bioelectromagnetics Society, Victoria, British Columbia, Canada, June 9-14, 1996. Session A-8: In Vivo Bioeffects, Abstract A-8-1. BEMS Abstract Book, p. 30.

Frey, A. H., 1993. Electromagnetic field interactions with biological systems. FASEB J. 7:272-281.

James, W. H., 1995. Sex ratio of offspring of female physiotherapists exposed to low-level high-frequency electromagnetic radiation. Scan. J. Work Environ. Health, 21:68-69.

Gubéran, E., Campana, A., Faval, P., Gubéran, M., Sweetnam, P. M., Tuyn, J. WN., and Usel, M., 1994. Gender ratio of offspring and exposure to shortwave radiation among female physiotherapists. Scand. J. Work Environ. Health, 20:345-348.

Lamont, J. W., Weber, R. J., Dooley, M. P., Pineda, M. H., and Moye, D. J., 1994. Shielded Culture chamber and controlled uniaxial magnetic field generator for very low frequency (VLF) magnetic field exposure of cells during **in vitro** culture. Presented at: Experimental Biology 94, Anaheim, CA., April 24-28, 1994, The FASEB Journal, Vol. 8, No. 4, Part 1, p. A398, March 1994.

Larsen, A. I., Olsen, J., and Svane, O., 1991. Gender-specific reproductive outcome and exposure to high-frequency electromagnetic radiation among physiotherapists. Scand. J. Work Environ Health, 17:324-329.

Zusman, I., Yaffe, P., Pinus, H., and Ornoy, A., 1990. Effects of pulsing electromagnetic fields on the prenatal and postnatal development in mice and rats: **In vivo** and **in vitro** studies. Teratology 42:157-170.

A DESCRIPTION OF INTEGRATED JOINT USE
INITIATIVES TO SATISFY CUSTOMER REQUIREMENTS
ACROSS GOVERNMENT, ACADEMIA, AND INDUSTRY

Andrew E. Jackson, Ph.D.
Assistant Professor
Department of Aeronautical Technology

Arizona State University
School of Technology
6001 South Power Road - SIM Building 425
Mesa, Arizona 85206

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

August 1996

17-1

A DESCRIPTION OF INTEGRATED JOINT USE INITIATIVES TO SATISFY CUSTOMER REQUIREMENTS ACROSS GOVERNMENT, ACADEMIA, AND INDUSTRY

Andrew E. Jackson, Ph.D.
Assistant Professor
Department of Aeronautical Technology
Arizona State University

Abstract

The joint use of facilities, equipment, and personnel resources has become a significant issue in recent years as a result of declining budgetary resources, downsizing initiatives, streamlining business operations, and reductions in operational forces. This pattern is evident in the military services, in national, state, and local governmental agencies, in academia, and in industry alike. The need to optimize limited resources, to minimize procurement costs, operational and maintenance (O&M) costs, and reduce duplication of effort has provided the impetus to identify resources that can be shared, based on mutually beneficial organizational objectives. The Armstrong Laboratory, Human Resources Directorate - Aircrew Training Research Division and Arizona State University have joined forces to address joint-development initiatives pursuant to Title 10 United States Code Section 2194 in which Education Partnership Agreements between governmental agencies and educational institutions are authorized and encouraged.

A DESCRIPTION OF INTEGRATED JOINT USE INITIATIVES TO SATISFY CUSTOMER REQUIREMENTS ACROSS GOVERNMENT, ACADEMIA, AND INDUSTRY

Andrew E. Jackson, Ph.D.

Introduction

The objective of any successful business is to optimize resources under its control while providing customers with high quality products at affordable costs. In order to achieve these objectives, an organization must optimize their performance by all available means. One such optimization technique involves the joint use of personnel, equipment resources and facilities. By reducing or eliminating resource duplication, an organization can minimize expenses in many different areas. A secondary organizational benefit of sharing resources is realized when utilization rates are improved for costly equipment and facilities in conjunction with highly trained and experienced personnel. These improved resource utilization rates define the most significant benefit of shared resource management.

This paper will address a few of the initiatives currently underway between the Armstrong Laboratory, Human Resources Directorate - Aircrew Training Research Division (AL-HRA) and Arizona State University (ASU) to enhance mutually beneficial resource utilization for both organizations. In response to Title 10, Subtitle A, Part III, Chapter 111 United States Code Section 2194, an Educational Partnership is being developed between AL-HRA and ASU to explore joint aviation education, research, and training opportunities. The research effort during

the Summer 1996 Air Force Office of Scientific Research - Summer Research Program (AFOSR-SRP) assignment has resulted in a preliminary Educational Partnership agreement between ASU and AL-HRA which is expected to be finalized during the last quarter of calendar year 1996.

Discussion

Asset utilization is a primary task for management and operational personnel at all levels of the organization. If a manager can reduce the cost of producing his/her component of an overall product or service, then it follows that cumulative costs will likewise be reduced, unless one or more groups achieve their unit efficiencies at the expense of other organizational units. If this condition occurs, however, it is the responsibility of upper management to resolve asset utilization conflicts so overall resource utilization is optimized. Assuming no resource conflicts exists then the organization may experience a condition where excess capacity is available. This excess capacity can also be viewed as an under-utilization of existing assets. These under-utilized assets form the basis for joint use and/or cooperative development initiatives.

Most organizations face the problem of managing under-utilized resources at one time or another due to cyclical variations in the customer's demand curves and corresponding funding cycles. As excess capacity remains in an organization, additional costs accrue. These accrued costs are similar to holding costs associated with on-hand inventory in a manufacturing firm. To reduce these *holding costs* and achieve more efficient utilization of resources, each participating unit can

benefit from shared resource utilization instead of developing independent capabilities to perform complementary or supplementary tasks.

Each shared resource reduces the need for new or stand-alone equipment, facilities, or personnel to support overall organizational goals and objectives. Under a shared resource model, joint utilization simultaneously reduces excess capacity and helps to reduce the cyclical variations in specific supply/demand curves based on fiscal, personnel, and/or procurement constraints. In order to determine which resources are the best candidates for cooperative use, each participating organization must identify their individual strengths and weaknesses in a specified area. After these individual and organizational strengths and weaknesses have been identified, the participating team members can negotiate for the resources needed to establish a complete, project-oriented approach to a specified problem. In so doing, the team will supplement each other's weakness in the overall project. The staffing analysis described above is similar to a traditional internal staffing or resource utilization analysis. The primary difference between an internal staffing analysis and the staffing techniques identified herein is the cross-organizational cooperation and trust which is needed to support and complete the defined tasks. The cross-organizational team approach takes on a different focus when independent organizations are used to define the project team.

Once project resources have been identified using the multi-functional/cross-organizational team model, the team leaders must be able to track team progress toward project and task completion.

Several techniques exist to monitor the status and resource utilization for a specified project or groups of projects. Some of these techniques include:

- Project Evaluation and Review Technique (PERT)

- Resource Leveling

- Constrained Resource Scheduling

- Heuristic Methods

- Optimizing Methods

- Multiproject Scheduling and Resource Allocation

- Mathematical Programming (Meredith & Mantel, 1995)

The techniques cited by Meredith & Mantel above define a small portion of the traditional project management tool set. Detailed descriptions of these techniques (and more) may be found in various project management references including: Badiru & Pulat (1995), Cook & Russell (1989), Obradovitch & Stephanou (1990), Salvendy (1982), and Sink (1985). Detailed descriptions of these project management, optimization, resource management, and scheduling techniques, however, are beyond the scope of this paper.

Kerzner (1995) has defined a mature organizational structure as one that has a relatively flat managerial infrastructure, and one that is complemented by non-dedicated project teams, informal project management, and a less critical reporting levels for project managers compared to a more traditional, hierarchical project management approach. The shared resource model is consistent with a mature organizational structure cited by Kerzner. It is important to note that a Kerzner's mature organizational structure is also consistent with the recent trend in industry to flatten the

organization's management structure through employee empowerment and downsizing (or right-sizing) to streamline business operations. Again, the concept of sharing resources within this model supports a more efficient use of personnel, facilities and equipment to satisfy specific project needs.

As organizations identify resource areas in which their inherent capabilities are inadequate to meet all specified project requirements, the manager has a few options as to how to meet the specified need. One option is to disregard the requirement and complete only that portion of the assigned task which he/she has the capability to complete. This approach clearly has limitations if the manager expects to receive repeat business from the specified customer. A second alternative is to procure additional equipment, to build new facilities, to hire new people, and solicit additional funds from the customer to support the project requirements. While this approach may enhance the manager's realm of control over organizational resources, it may prove to be too costly and inefficient in the long run. A third alternative is to identify cooperative organizations who have the type and number of resources needed to augment specific project requirements. As cooperative agreements are forged between partnership organizations, the immediate needs of the project can be satisfied and a stronger strategic alliance developed to respond to future requirements. For example, if one organizational unit has a need for an expensive piece of test or support equipment, it is likely other units in the organization (or on the team) will also benefit from using that same equipment on a time-share basis, rather than procuring their own units which could remain idle for large blocks of time. Since sharing an asset reduces the overall costs to the organization (or more generically, to the project effort itself) then the cost to benefit ratio is

17-7

enhanced proportionately. Additionally, future project requirements can be reinforced by establishing long-range cooperation between organizations who share common objectives but who do not compete for common funding sources and customers.

Methodology

To facilitate the cooperative training and research initiatives cited above, a preliminary cooperative agreement was drafted and circulated for comment among key members of the participating organizations (AL-HRA and ASU). Following a brief review period, comments from both organizations were compiled into a working draft for distribution to outside agencies for further review and comment. The draft version of the Educational Partnership Agreement (Enclosure (1)) was developed as an umbrella agreement under which current and future educational, research, and training initiatives would be structured. By establishing a permanent working relationship between neighboring organizations who share common research and education objectives, a synergistic bond can be realized which will benefit all participating groups.

Results

The AFOSR - SRP assignment has enabled Arizona State University to move closer to establishing a long-term educational, research, and training partnership with the Armstrong

Laboratory in Mesa, Arizona. The opportunity to meet and work with the scientists and engineers at the Armstrong Laboratory provided an invaluable insight into current and projected research requirements. The Educational Partnership Agreement which was developed jointly by ASU and AL-HRA will facilitate future cooperative projects where common research goals can be established. The Educational Partnership Agreement is currently under review by both participating organizations and is expected to be finalized by the end of Calendar Year 1996 (CY-96). Once the agreement is finalized, research partnerships can be established between individual scientists and engineers across the organizations, enabling the participants to strengthen their current research capabilities and to build a world-class aviation research center at the William Campus.

Conclusion

The need to improve resource allocations and optimize investment dollars at the Williams Campus has led to the development of a preliminary agreement between the Armstrong Laboratory (AL-HRA) and Arizona State University to share research and training assets in order to optimize limited resources from each organization. The combined AL-HRA/ASU research team has agreed to collaborate on joint training, research, and development opportunities at the Williams Campus. This collaboration will extend to joint use of facilities, personnel, equipment, and support systems on a non-interference basis. Research support from Arizona State University will come in the form of student research assistants, faculty advisors and scientists, laboratories, and

aviation support personnel. The AL-HRA contribution will include research task definitions through externally funded projects from various sponsors, including the U. S. Air Force, the State of Arizona, the Federal Aviation Administration (FAA), and a variety of local, regional and national organizations, facilities, equipment, research personnel, and research topics similar to those identified below:

- Night Vision Simulation Research

- Visual Psychophysics

- Instructional Simulation Development and Evaluation

- Information Technology

- Head and Eye Movement Behavior

- Instrumentation Research

- Virtual Reality Research

- FAA Approach Lighting Systems research

The AL-HRA and ASU team is working closely to build a new paradigm for education, training, and research through a cooperative, consortium-based management approach. This model has already proven useful on a limited scale but it is expected to serve as the foundation for several long-range programs at the Williams Campus.

EDUCATION PARTNERSHIP (AL-EP-9601) BETWEEN

ARMSTRONG LABORATORY and THE ARIZONA BOARD OF REGENTS

ACTING FOR ARIZONA STATE UNIVERSITY

I. Preamble

The Department of the Air Force, Armstrong Laboratory, Aircrew Training Research Division, Mesa, Arizona, and The Arizona Board of Regents acting for Arizona State University, Tempe, Arizona, (hereinafter referred to as "the Parties") hereby enter into this Education Partnership Agreement pursuant to Title 10 United States Code Section 2194.

The terms and conditions of this Agreement are set forth as follows.

II. Introduction

The Parties enter this Agreement in recognition of the importance of education and research to the future and economic well-being of the nation, as well as the importance of Armstrong Laboratory to business, industrial, and educational institutions in Arizona. Nothing in this agreement shall be deemed to create a formal business relationship of joint venture or partnership between the Parties until specific contracts are established under this agreement.

III. Objective

The objective of this Agreement is to facilitate cooperative utilization of resources (e.g. facilities, equipment, personnel, publications, and various research laboratories) between the Armstrong Laboratory and Arizona State University as part of the Education, Research and Training (ERT) Consortium located at the former Williams Air Force Base in Mesa, Arizona. Specific objectives will be negotiated independently and incorporated into this agreement as required by sequential attachments to this signed document. The Armstrong Laboratory and Arizona State University plan to conduct research in the following areas, leading to cooperative research opportunities and enhanced operational capabilities for both of the Parties.

Research areas may include (but shall not be limited by) the following:

- Education and Training Research and Development

- Improvements in Human Factors Design and Implementation

- Enhancement of Human Performance

- Aircraft Systems Development and Evaluation

- Aircraft Performance Evaluation

- Aviation Navigation and Guidance Systems

- Computer Science (including Software Development and High Speed Processing)

- Behavioral Sciences (including Vision, Perception, and Cognition Research)

- Statistical Research and Experimental Design

- Visual Systems Capabilities and Limitations

- Simulation and Simulation Support Technologies

- Technology Integration and Technology Transfer Issues

- Joint Use and Cooperative Development

Access to shared resources (e.g., hardware, personnel, facilities, etc.) will provide a unique opportunity for university students, academic professionals, and faculty to work on research projects which would otherwise be unavailable in an educational environment.

IV. Partnership Definition

Pursuant to statutory authority under Title 10 United States Code Section 2194, Armstrong Laboratory is authorized to participate in educational partnerships for the purpose of encouraging and enhancing study and research in scientific disciplines. Specific support elements for the specified educational partnership are cited in the aforementioned statute.

V. Specific Obligations

    A. Arizona State University shall: --

        (1) Be responsible for all ASU faculty, ASU student, and other ASU employee salaries or other compensations and benefits as necessary (except funded internships/co-operative educational programs).

Enclosure (1)

(2) Provide equipment and adequate ASU facilities to support research objectives, including library access for partnership participants;

(3) Assemble academic, industry, private, and governmental experts to support research opportunities;

(4) Survey available experience of ASU faculty and students to support joint research objectives;

(5) Develop and implement research proposals through joint cooperation with Armstrong Laboratory scientists, managers, and engineers;

(6) Test and evaluate new training and operational technologies, methods, applications, and techniques as defined in joint research proposals;

(7) Provide Armstrong Laboratory with a copy of any significant findings that are documented as a result of activities for items in paragraphs V. A. (2) through V. A. (6) above.

B. Armstrong Laboratory shall:

(1) Make available hardware, facilities, personnel, etc.

(2) Provide three office spaces for use as follows:

    a. Two ASU faculty offices

Enclosure (1)

     b. One research assistant (student) office;

(3) Provide equipment and adequate facilities to support research objectives, including government library access for partnership participants;

(4) Make laboratory personnel available to teach or assist in the development of programs and/or courses for Arizona State University;

(5) Cooperate with Arizona State University in developing a program under which students may be given academic credit for work on laboratory research projects;

(6) Provide academic and career advice and assistance to students of Arizona State University;

(7) Support funded and unfunded graduate and undergraduate internships and co-operative educational opportunities.

(8) Offer currently existing experienced personnel for consultation and joint proposal development (including support contractors to Armstrong Laboratory);

(9) Grant Armstrong Laboratory secure facility access to selected ASU faculty members to facilitate student and faculty research objectives;

(10) Assist in identifying academic, industry, private, and governmental experts to support research opportunities.

C. Subject to Article VIII (Risk of Loss) below, neither party shall be obligated to compensate the other party for costs incurred by the other party in carrying out activities deemed appropriate under this partnership agreement. Facility impact issues will be negotiated in good faith to seek equitable trade-off capabilities during the term of this Agreement.

D. Each party shall direct its own activities pursuant to this partnership agreement. Neither party shall have authority to direct the other's activities.

VI. Benefits

A. The benefits to Arizona State University include:

(1) A formal vehicle for information exchange with Armstrong Laboratory's Aircrew Training Research Division, the Air Force's premier organization for research and development in aircrew training techniques and technologies. The division's basic mission is to increase aircrew effectiveness through enhanced training.

(2) Access to state-of-the-art equipment and technology that would not normally be available to the university community.

(3) Opportunities for research students and faculty to understand and improve upon technology areas that have been jointly developed by government, academia, and industry.

B. The benefits to Armstrong Laboratory include:

(1) Promoting the education of future scientists and engineers;

(2) A new perspective on projects by bringing in new ideas from academia;

(3) Enhancing Armstrong Laboratory scientists and engineers by exposure to faculty, academic professionals and dedicated students working in relevant areas of research.

(4) Joint use of facilities, personnel, equipment, publications, and expertise unavailable to Armstrong Laboratory.

VII. Partnership Administration

The administration of this Agreement and the coordination of specific activities that comprise this program will be the joint responsibility of the designated program managers from Armstrong Laboratory and Arizona State University.

One or more designated individuals will serve as the partnership program manager(s) on behalf of Armstrong Laboratory. These individuals will work with the program managers from Arizona State University to identify, select, and prioritize the activities in which the Parties engage pursuant to this Agreement and will ensure that program activities meet the statutory and regulatory requirements of the Federal Government and the Department of the Air Force.

Designated individuals from Arizona State University will serve as the partnership program managers on behalf of Arizona State University Main Campus and East Campus respectively.

They will work with the program manager(s) and research engineer(s) from Armstrong Laboratory to identity, select, and prioritize activities in which the Parties engage pursuant to this Agreement and will ensure that program activities meet the statutory and regulatory requirements of Arizona State University.

VIII. Risk of Loss

Arizona State University may inspect Armstrong Laboratory property provided pursuant to this Agreement prior to use. Such property may be repaired or modified at Arizona State University's expense only after obtaining the written approval of Armstrong Laboratory. Any repair or modification of the loaned property shall not affect the title of Armstrong Laboratory. Unless Armstrong Laboratory hereafter otherwise agrees, Arizona State University shall, at no expense to Armstrong Laboratory, return all unmodified, loaned property to Armstrong Laboratory after termination or expiration of this Agreement in the condition in which it was received (normal wear and tear excepted). If the loaned Armstrong Laboratory equipment was approved for modification during the period of ASU custody, the equipment may be returned to Armstrong Laboratory as modified rather than converting the equipment to an original, unmodified condition prior to return (normal wear and tear excepted).

IX. Period of Agreement

The term of this Agreement is for a period of 36 months, commencing on the date of the last signature affixed below. Either Party may terminate this Agreement earlier upon delivery of

written notice at least thirty (30) days in advance. If either Party requests modification of this

Agreement, the Parties shall, upon reasonable notice of the proposed modification by the Party

desiring the change, confer in good faith to determine the feasibility of such modification.

Modifications shall not be effective until a written amendment is signed by duly authorized

representatives of the Parties.

X. Notice of State Statutes

Notice is hereby provided of Arizona Revised Statutes, sections 12-133, 12-1518, and 38-511.

IN WITNESS WHEREOF, the Parties have caused this Agreement to be executed in duplicate.

ARIZONA BOARD OF REGENTS

FOR ARIZONA STATE UNIVERSITY: ARMSTRONG LABORATORY:

By: _____   By: _____

(Signature)   (Signature)

Name: _____   Name: _____

Title: _____   Title: _____

Date: _____   Date: _____

Enclosure (1)

References

Badiru, A. and Pulat, P. (1995), Comprehensive project management: Integrating optimization models, management principles, and computers, Englewood Cliffs: Prentice Hall, Inc.

Cook, T. and Russell, R., (1989), Introduction to management science, Englewood Cliffs: Prentice Hall, Inc.

Kerzner, H. (1995) Project management: A systems approach to planning, scheduling and controlling, New York: Van Nostrand Reinhold.

Meredith, J and Mantel, S. Jr. (1995), Project management: A managerial approach, Third edition, New York: John Wiley & Sons, Inc.

Obradovitch, M. and Stephanou, S. (1990), Project management: Risks & productivity, Bend, Oregon: Daniel Spencer Publishers.

Salvendy, G.(1982) Handbook of industrial engineering, New York: John Wiley & Sons, Inc.

Sink, D.(1985). Productivity management: Planning, measurement and evaluation, control and improvement, New York: John Wiley & Sons.

United States Code. Title 10, Subtitle A, Part III, Chapter 111, Section 2194. Educational Partnerships.

John Kalns report was not available at the time of publication.

MODELING DECOMPRESSION SICKNESS USING
SURVIVAL ANALYSIS TECHNIQUES

Nandini Kannan
Assistant Professor
Division of Mathematics and Statistics

Aparna Raychaudhuri
Assistant Professor
Division of Mathematics and Statistics

University of Texas at San Antonio
6900 N. Loop 1604 West
San Antonio, TX 78249

MODELING DECOMPRESSION SICKNESS USING
SURVIVAL ANALYSIS TECHNIQUES

Nandini Kannan
Assistant Professor
Division of Mathematics and Statistics
University of Texas at San Antonio

Abstract

The application of survival analysis methods to modeling altitude decompression sickness was studied. Parametric Models were developed to assess the effect of several covariates on the probability of DCS. The model was used to predict the risk/ probability of DCS for several flight profiles. The goodness of fit of the model was examined using cross validation techniques. These models were adjusted to include data on bubble grade and times.

MODELING DECOMPRESSION SICKNESS USING
SURVIVAL ANALYSIS TECHNIQUES

Nandini Kannan

Aparna Raychaudhuri

Introduction

Decompression sickness (DCS) is caused by exposure to significant change in environmental pressure, encountered during diving, high altitude exposures or artificially induced pressure changes in hyperbaric or hypobaric chambers. Researchers in the field believe that DCS occurs as a result of the growth of bubbles in body tissue. For large and rapid pressure reductions, supersaturation occurs as a result of the inability of tissue gas exchange processes to expel excess nitrogen. These gases which come out of solution when tissues are sufficiently supersaturated collect as bubbles in the tissue. The size and location of these bubbles are thought to have a significant effect on the resulting DCS symptoms.

It is also well known that the incidence of DCS depends on the length of exposure, final altitude, level of physical activity, and amount of prebreathing time. It is possible to decrease/ eliminate the risk of DCS by controlling these so called "risk factors". Air Force personnel are routinely exposed to high altitudes, and an accurate assessment of DCS risks is vital to a successful mission. The High Altitude Protection Function, Brooks Air Force Base is frequently called upon to assess risks for both civilian and military personnel under specified flight protocols. To provide answers to these questions, and also to obtain a clearer understanding of the effects of various risk factors, the AL/ CFTS group has been working on an approporiate model to predict DCS risk using physical and physiological principles.

For the past several years, experiments have been conducted in a hypobaric chamber at the Armstrong Laboratory using human subjects. The subjects were exposed

to different altitudes, varying denitrogenation times, and different prebreathing mixtures. Several measurements were recorded during the experiment including onset time of DCS, physical activity, time spent in the chamber etc. The data collected includes a measure of bubble grade measured on a Spencer scale using echo cardiography.

Background

The literature in DCS modeling has mainly focussed on diving experiments. However, there are certain characteristics that are unique to high altitude exposure. One significant difference is the faster ascent rate in high altitude situations compared to the slower ascent rates for divers. There is also a significant threat of instantaneous pressure loss in aircraft.

Altitude DCS modeling has primarily dealt with bubble growth models. Van Liew et al [1] developed a probabilistic model of altitude DCS. They assumed that the risk of DCS is related to the number of bubbles and the volume of gas that can be liberated from a unit of tissue. Mathematical equations were formulated in order to explain these phenomena. Several models were tested to examine the effects of duration of 100% O_2 at ground level (prebreathing), atmospheric pressure after ascent, and exposure duration.

Gerth and Vann [2] developed an extensive model for bubble dynamics to provide an assessment of DCS. The bubble dynamic equations were similar to those in Van Liew et al. The percentage of individuals with DCS was considered to be the response variable, and maximum likelihood methods used to estimate the model parameters. The authors discussed the need for onset times to be included to develop more realistic models.

The models based on bubble growth do have some drawbacks. First, several assumptions have to be made in order for these equations to hold. There is no data available on many parameters that are used in these models. These models also

ignore the time to onset of DCS and provide limited insight into the effects of different covariates on the risk of DCS.

Survival analysis methods seem to be most appropriate for modeling DCS risk, because the time to occurrence of DCS is a random event. It is also known that not all individuals exposed to identical conditions exhibit DCS symptoms. Kumar et al [3], [4], [5] in a series of papers, discussed survival models for predicting DCS. They developed logistic and loglinear models to predict DCS as a function of Tissue Ratio which is a measure of tissue nitrogen decompression stress, and CMB (circulating microbubbles) status. The models used the logarithm of time to DCS and maximum likelihood techniques to estimate the model parameters. In a recent article, Conkin et al [6] used loglogistic survial models on the data from 66 NASA and USAF hypobaric chamber tests. The models examined the effects of TR and exercise status (yes, no) on the risk of DCS.

In this report, we develop a survival model for DCS risk using several risk factors. The model will be used to predict the probability of DCS over time for different flight protocols. The model will be adjusted to also include information on bubble times and grades.

In the next section, we discuss the application of survival analysis to this problem.

Methodology

Survival Analysis refers to the study of lifetime or failure time of subjects or components. The primary variable of interest in this study is the time to onset of DCS for an individual. An important problem is to identify the risk factors (also referred to as covariates) that have a significant effect on the incidence of DCS. Some important definitions and terminology from survival analysis are included in the appendix. For a detailed discussion, the reader is referred to Lawless [7] and Lee [8].

The survival time T is modeled using various forms of the probability density func-

tion (pdf). The common models include the exponential, gamma, Weibull, lognormal, and loglogistic. Each pdf has its unique functional form, and defines a corresponding survival function and cdf.

The risk function changes according to the underlying survival distribution and can take different shapes. From empirical evidence, it is known that the risk of DCS initially increases upto a certain time point, and then decreases because of denitrogenation. The loglogistic and lognormal distributions have risk functions that fall into this category.

A characteristic unique to survival data is "censoring" or incomplete information. The instant a subject experiences any symptoms, the experiment is aborted and the subject brought down to ground level. However, individuals who do not report any symptoms will remain in the chamber for the duration of the flight. If the flight duration were prolonged, these subjects could possibly exhibit symptoms: these are censored observations. For individuals who experience DCS, the time to onset of DCS is available. For all other subjects, the time to DCS is considered to be beyond the duration of the experiment.

Once we select an appropriate model, the likelihood function can be written. This function describes the probability of observing the actual data points and includes information from both censored and uncensored individuals. The likelihood is maximized to obtain estimates of the underlying parameters. In the next section, we fit survival models to the USAF data and illustrate the fit of the models.

Results

The data collected in the study included the time to onset of DCS symptoms (ONSET), time of prebreathing (BR), pressure (PRES), time at maximum altitude (TALT), and exercise code (EX). Different types of exercise were performed by the subjects. They were classified according to the amount of oxygen consumption as rest, mild, and heavy.

The dataset contained several profiles with pressures above 314 mm. In this range, the rate of DCS is extremely small. There were only 12 cases of DCS among 451 subjects. The DCS rate for lower pressures was significantly higher. It was decided to use only the 975 datapoints (with pressures less than or equal to 314) for the statistical analysis. Similar models could be constructed for the complete dataset.

For individuals reporting no symptoms, the onset time was replaced by their corresponding TALT times. These are the censored observations. A censoring variable (CENSOR) was created to indicate the status of each individual: 1–DCS and 0–No DCS. The data set contained 469 censored observations (48 % DCS incidence).

Prebreathing times are known to have a significant effect on reducing DCS risk. However, a cursory examination of the database indicated an increase in the rate of DCS for increased prebreathing times. This anamoly was because higher prebreathing times are commonly observed for flight profiles with long exposure times and higher altitudes. To adjust for these differences, a new variable BRTALT was created which is the percentage of total exposure time spent prebreathing. Clearly a high value of BRTALT indicates one of two possibilities: longer prebreathing times or shorter flight durations. High values of this variable are associated with low risks of DCS which is in line with empirical evidence.

The loglogistic and lognormal distributions were fitted to the data, with maximum likelihood techniques used to obtain the parameter estimates. Three covariates or risk factors were included in the model: PRES, BRTALT, EX. The pressure values were included on a log scale. Weights were incorporated into the likelihood equation to adjust for the large dispersions in the data set.

The output from the SAS programs (Table 1) provides the parameter estimates, the variance of the estimates and a chi-square value used to determine the relative importance of the different factors. Pressure was clearly the most important determinant, followed by exercise and the proportion of total exposure time spent in prebreathing. The coefficients for all three covariates were highly significant (p-value

Table 1: Parameter estimates for the weighted model

Variable	DF	Estimate	Std. Err.	Chi-sq	p-value
INT	1	-8.00	2.45	10.63	0.0011
PRES	1	2.53	0.44	32.57	0.0001
BRTALT	1	1.29	0.39	11.26	0.0008
EX	1	-0.53	0.14	13.68	0.0002
SCALE	1	0.60	0.03		

< 0.001), indicating their importance on the risk of DCS. The loglikelihood value (a measure of how well the model fits the data) was -560.84.

A lognormal model was also fit to the data, and the loglikelihoods of the two compared. The predictions from both models were almost identical, providing very little evidence of the superiority of any one distribution. It was therefore decided to use the loglogistic model in all further analyses because of its' simpler form. See the appendix for the form of the survival and risk functions.

Interaction terms included in the model did not provide much improvement. In some cases the interactions tended to mask the effects of the main factors. This confirmed the belief that these three primary risk factors were sufficient to adequately predict the probability/ risk of DCS.

Since the main focus of the study is to develop a model that predicts accurately, we need to validate the model. The first step in this exploratory process was to obtain the estimated cdf for two flight profiles with PRES = 231, EX = Mild, TALT = 240, and BR = 135, 75. All individuals in the profile with BR = 135 were deleted from the database, and the loglogistic model fit to the remainder of the data points. An overlay of the estimated cdf from this truncated dataset and the observed onset times provides a cross validation of the model (Figure 1). There is a very close agreement
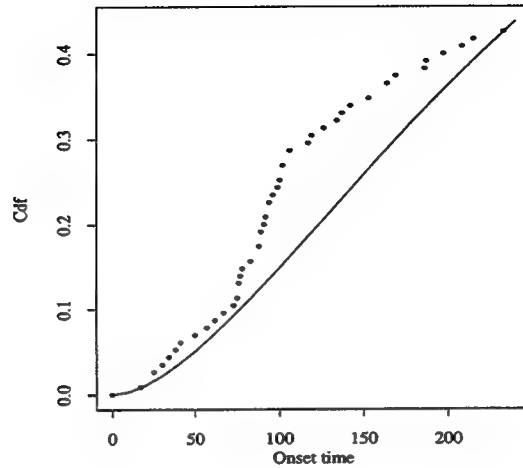
Figure 1: Cross Validation: BR = 135

between the estimated probability of DCS and the observed cdf. This trend was visible for several other profiles as well.

Figure 2 shows the predicted and observed cdf's for the profile with BR = 75. This predicted model is caluclated from the entire dataset to examine how well the model fits the underlying data. Based on these observations, this model, though simple, provides extremely good predictions.

These validations and cross validations confirm our beliefs that this model does indeed describe the nature of DCS and can provide very accurate predictions. We would now like to determine the effects of different covariates on the DCS probability/ risk.

To examine the effects of prebreathing times and exercise on DCS risk or probability, we provide some graphs. Figure 3 examines the effects of exercise on the probability of DCS. As exercise levels move from rest to mild to heavy, the probability of DCS increases by approximately the same amount. Figure 4 provides the corresponding risk functions for the three prebreathing times 0, 60, and 90 minutes.
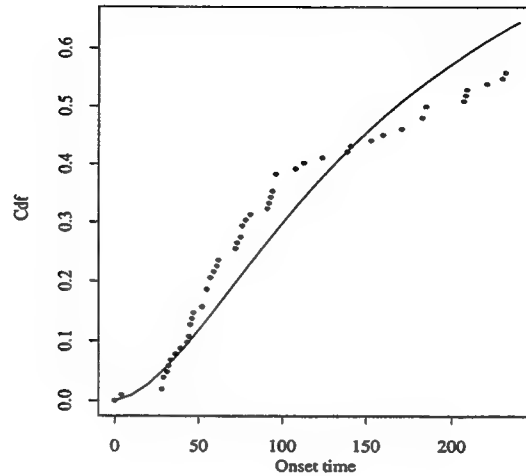
Figure 2: Validation: BR = 75

The risk is clearly highest for 0 prebreathing, the difference in the risks for 60 and 90 minutes are not that large.

Loglogistic Models for Time of onset of bubbles

As mentioned in the introduction, one of the primary causes of DCS is thought to be the size and location of bubbles in body tissue. The experiments in the hypobaric chamber included measurements on bubbles using a Doppler ultrasound. Both audio and video images provided information on bubble density of bubbles recorded at fifteen minute intervals. The bubbles were graded on a Spencer scale with values ranging from 1 (low) to 4 (high). From this data, the time at which the bubble grade is maximum was extracted along with the corresponding grade. The subjective nature of bubble grade data, led to the following classification for the statistical analysis. Bubbles were classified as grade low (grades 0,1) or high (grades 2,3 and 4).

Since bubble grades were classified as a binary response, a logistic model was fit to the probability of low and high grades. This probability depends on the time of
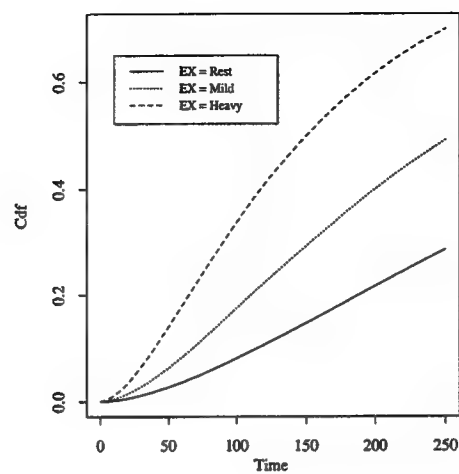
Figure 3: Predicted Cdfs: PRES = 282, TALT = 240, EX = Mild
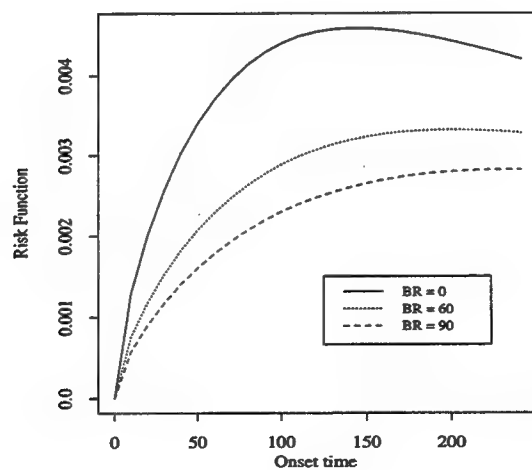
Figure 4: Predicted Risk Function: PRES = 282, TALT = 240, EX = Mild

19–11

exposure, prebreathing time, and exercise status. In the presence of these covariates, pressure was found to have no significant effect. The logistic model helped to identify those profiles for which high grades were more likely to occur compared to low grades. In general, it was found that subjects who performed mild exercise in flights of moderate/long duration (240 minutes or longer) were more likely to have high bubble grades (i.e. probability of observing high grades exceeded 50 %). For individuals at rest with low prebreathing times, high grades were more likely to occur. If prebreathing times were moderate or high, the individuals almost always had very low bubble grades.

The logistic model provides information about those profiles for which bubble data would have a significant effect on the predictions. The next step in assimilating the bubble data into the DCS model is to examine the times at which bubbles were observed.

The time at which the maximum grade is attained could be interpreted as a 'survival time' and modeled accordingly. Clearly this time is affected by the amount of prebreathing time, exercise status, and pressure. The risk function increases steadily, reaches a maximum, and then decreases because of denitrogenation. Once again, loglogistic and lognormal models are found to be appropriate. The models are fit to only those individuals who had high bubble grades. The loglogistic model provides information on the chance of observing high bubble grades within a specified time period as a function of several risk factors, and can be used to determine the percentiles for MAXT. In particular, the median will represent the time by which 50 % of individuals will have high bubble grades and thus provide an "estimate" of the time at which the bubble grade is maximum. Tables 2 and 3 provide the 25th, 50th, and 75th percentiles for two flight profiles. From the first table, we observe that 25 % of individuals with no prebreathing will have high bubble grades within 47 minutes, this interval increases to 69 minutes with 60 minutes of prebreathing. The effect of pressure/ altitude is quite pronounced. When the pressure drops to 226

19–12

Table 2: Percentiles of MAXT: PRES = 282, and EX = 2 (Mild)

BR	Percentile		
	25	50	75
0	46.7	69.4	103.0
60	69.3	102.9	152.9
90	72.0	106.9	158.9

Table 3: Percentiles of MAXT: PRES = 226, EX = 2 (Mild)

BR	Percentile		
	25	50	75
0	36.7	54.5	80.9
60	54.5	80.9	120.1
90	56.6	84.1	124.8

with 0 prebreathing, 50 % of individuals have high bubble grades within 55 minutes compared to 70 minutes for a pressure of 282. There is almost a 20 minute difference in the percentiles for the two pressures.

This information on bubble times is significant and should provide more accurate predictions of the probability of DCS over time. The question remains: How do we use this information in a DCS model? In order to provide an answer, the covariate MAXT was added to the loglogistic model. The EX covariate was found to be nonsignificant in this model. However, exercise is still a part of the predictions because it has an effect on MAXT. The model was fit for the two grade categories:high and low. Fow low bubble grades, as expected the effect of MAXT was negligible. However, from the

Table 4: Parameter estimates for the weighted model

Variable	DF	Estimate	Std. Err.	Chi-sq	p-value
INT	1	-3.66	1.80	4.12	0.0424
PRES	1	1.34	0.31	17.37	0.0001
BRTALT	1	0.96	0.30	10.21	0.0014
MAXT	1	0.01	0.00	183.58	0.0001
SCALE	1	0.37	0.02		

table given below, it is clear that MAXT has a tremendous effect on the risk of DCS for high bubble grades. The time to onset of high bubble grades provides almost all the answers to predicting the chance of DCS. The influence of MAXT dampens the effect of all the other covariates. Since, the primary focus of this investigation is prediction of DCS for varying flight profiles, input values for the MAXT variable are necessary. One option is to use the median or 75th percentile as an input in the model. These values can be obtained from the loglogistic model fit to the bubble onset times. The 75th percentile provides a conservative estimate since only individuals with recorded bubbles are used in the model.

The graphs (Figures 5 and 6) show the effect MAXT has on the predictions. In the first graph, when MAXT is low (25th percentile), the curve rises very quickly. However, if the value of MAXT is high, the curve is less steep. The points scattered in the plot represent the observed probabilities. All three choices of MAXT provide very close estimates at 240 minutes. For the same profile, the model without bubbles predicted the final probability to be 60 %. There is a remarkable improvement with the addition of bubble data. Similar trends are observed for the other flight profile.

Conclusion
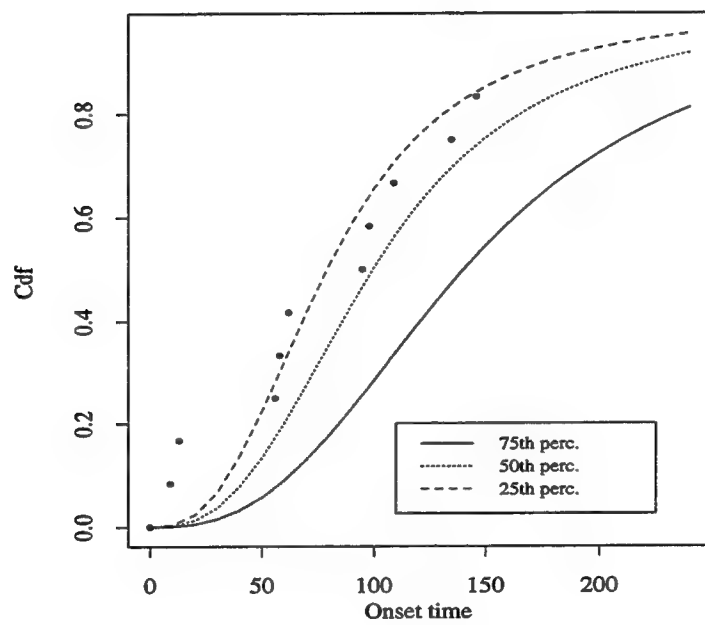
Figure 5: Predicted Cdf: PRES = 282, TALT = 240, EX = Mild, BR = 0
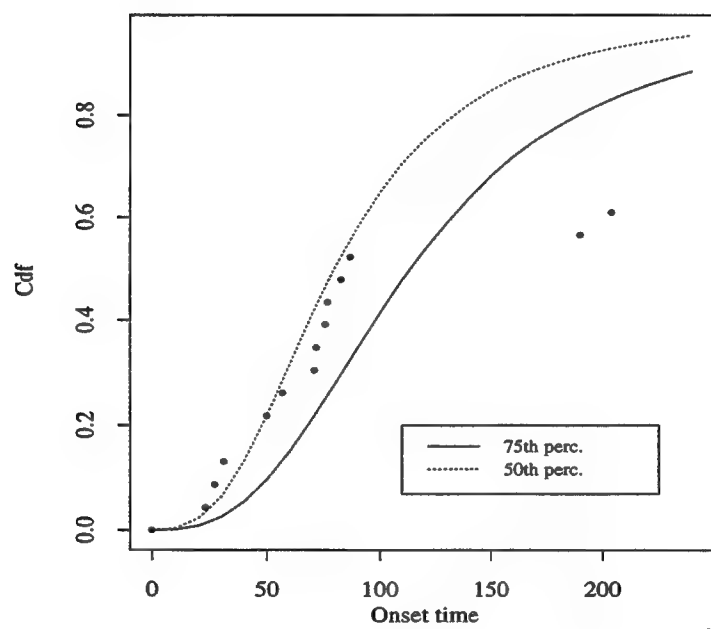
Figure 6: Predicted Cdf: PRES = 226, TALT = 240, EX = Mild, BR = 15

The loglogistic model with appropriate weights provides fairly accurate estimates of the probability of DCS over time, as a function of certain risk factors. The model also helped to identify ways of reducing the risk of DCS, by controlling the levels of the covariates. For example, longer prebreathing times will greatly diminish the risk of DCS over time. The cross validations provided strong evidence of the models' predictive abilities.

The bubble data was assimilated into the loglogistic model and clearly showed dramatic improvements especially for flight profiles with low prebreathing times.

The bubble growth model being developed at the Armstrong Laboratory will play a significant role in the ultimate DCS model. The bubble growth model will provide input on MAXT, and provide predictions for DCS risks. This approach of tying in the statistical and mathematical models should provide a long term solution for predicting/ assessing DCS risks.

Bibliography

H. D. Van Liew, J. Conkin, and M. E. Burkard, "Probabilistic model of altitude decompression sickness based on mechanistic premises," *Journal of Applied Physiology*, vol. 76, pp. 2726–2734, 1994.

W. A. Gerth and R. D. Vann, "Statistical bubble dynamics algorithms for assessment of altitude decompression sickness incidence," tech. rep., Armstrong Laboratory, 1995.

K. Kumar, B. Waligora, and D. Calkins, "Threshold altitude reuslting in decompression sickness," *Aviation, Space, and Environmental Medicine*, vol. 61, pp. 685–689, 1990.

K. Kumar, D. S. Calkins, J. Waligora, J. Gilbert III, and M. Powell, "Time to detection of circulating microbubbles as a risk factor for symptoms of altitude decompression sickness," *Aviation, Space, and Environmental Medicine*, vol. 63, pp. 961–964, 1992.

K. Kumar and M. Powell, "Survivorship models for estimating the risk of decompression sickness," *Aviation, Space, and Environmental Medicine*, vol. 65, pp. 661–665, 1994.

J. Conkin, K. Kumar, M. Powell, P. Foster, and J. Waligora, "A probabilistic model of hypobaric decompression sickness based on 66 chamber tests," *Aviation, Space, and Environmental Medicine*, vol. 67, pp. 001–008, 1996.

J. F. Lawless, *Statistical Methods for Survival Data Analysis*. Wiley, New York, 1982.

E. T. Lee, *Statistical Methods for Survival Data Analysis*. Wiley, New York, 1992.

Appendix

We introduce some of the common definitions and terminology used in survival analysis literature.

Let T denote the time to onset of DCS. The survival function is defined as

$$S(t) = P(T \geq t),$$

i.e. the probability that the individual is symptom free upto time t. The survival function is a nonincreasing (decreasing) function of time t with the properties that

$$S(0) = 1, \qquad S(\infty) = 0,$$

i.e. the probability of being symptom free for an infinite time is zero.

The Cumulative Distribution Function (cdf) is defined as

$$F(t) = P(T \leq t) = 1 - S(t),$$

the probability that DCS symptoms occur before time t. The Cdf starts at 0 and increases to 1. A steep cdf indicates the rate of survival is low or that onset time is relatively quick. A gradual, slow rising graph indicates the time to onset is higher. The survival function or cdf can be estimated from the data very easily using,

$$\hat{F}(t) = \frac{\text{\# of individuals with onset times less than } t}{\text{total \# of individuals}}.$$

Th probability density function (pdf) $f(t)$ of the survival/onset time T is defined as the probability that an individual exhibits DCS symptoms in a small interval per unit time. The usual relation between the cdf and the pdf is:

$$F(t) = \int_0^t f(x)dx.$$

The hazard or risk function $r(t)$ specifies the instantaneous rate of developing the symptoms at time t, given that the individual is symptom free up till t. Therefore $r(t)$ can be interpreted as the conditional failure rate. The risk function is appealing

since it describes the way in which instantaneous risk of DCS changes with time in the chamber. The risk function is defined as

$$r(t) = \frac{f(t)}{S(t)}$$

where $f(t)$ is the probability density function of the onset time as defined before.

For the loglogistic distribution, the survival and risk functions are

$$
\begin{aligned}
S(t) &= \frac{1}{1 + (\lambda * t)^{(\gamma)}} \\
r(t) &= \frac{\lambda\gamma(\lambda * t)^{(\gamma - 1)}}{1 + (\lambda * t)^{(\gamma)}}.
\end{aligned}
$$

The parameter λ depends on the covariate vector x,

$$\lambda = exp(-\beta' x)$$

where β is a vector of unknown parameters which will be estimated from the data. The parameter γ is a scale parameter.

The likelihood equation is given by

$$L(\beta) = \prod_{i=1}^{M} f(t_i) \prod_{j=1}^{N-M} F(t_j)$$

where M is the number of uncensored observations.

When weights are included, the likelihood is proportional to

$$L(\beta) = \prod_{i=1}^{M} [f(t_i)]^{w_i} \prod_{j=1}^{N-M} [F(t_j)]^{w_j}.$$

SKILL EVALUATION OF HUMAN OPERATORS

A. J. Koivo
Professor
Electrical and Computer Engineering

Purdue University
1285 Electrical Engineering Building
West Lafayette, IN 47907-1285

SKILL EVALUATION OF HUMAN OPERATORS

A. J. Koivo
Professor
Electrical and Computer Engineering
Purdue University

Abstract

The skills of humans performing tasks are often evaluated using Fitts' law, which is described by a straight line in coordinate system of execution time and task difficulty index. Then, the inverse of the slope of the straight line is called the execution capacity of the human.

The aforementioned variables are calculated from data of experiments which usually can be considered random variables due to human factors involved. In this framework, our study focuses in the variability of variables in Fitts' law and in the execution time. Indeed, we have determined the mean and standard directions for *individual* performances and also for the performances of a *group* of individuals. Thus, the skills of an individual can better be compared with those of the other individuals.

Fitts' law assumes that the performance demonstrating skills in a task is described by one variable. However, Fitts' law is ageometric approach which is difficult to apply when the task performance is characterized by two or more variables. Our study introduces a novel probabilistic skill index to evaluate skills in tasks; it chan be used even when a task is characterized by two or more variables. The probabilistic skill index does provide a broader basis for the skill evaluations of individuals. Moreover, the individuals performing tasks requiring certain skills can then be compared objectively on the basis of the probabilistic skill index.

SKILL EVALUATION OF HUMAN OPERATORS

A. J. Koivo

1. INTRODUCTION

When a person performs a task, the success of the performance partly depends on the difficulty of the task. A possible way to quantify the complexity of a task is by specifying an index of difficulty (ID), as proposed in [1]. The ability of a person to perform a task is often referred to as the skill that he/she has acquired. A person often gains skills by using his/her past knowledge and experience and by repeating the task performance. The levels of skills are difficult to evaluate and quantify for many tasks. Yet, the skills of a person in some applications can be quantified by evaluating his/her capacity to execute a specific task. This concept, "the execution capacity", is analogous to the channel transmission capacity introduced by Shannon in Communication Theory. It is determined on the basis of one variable which can be measured in the experiments, for example the execution time, or it is calculated from experimental data. When such a variable is graphed as a function of the difficulty index of the task, a straight line can be fitted to the data points, for example, by the regression analysis. This straight line represents so-called Fitts' law. The inverse of the slope coefficient in the straight line equation is the execution capacity of the person performing the task. Thus, the skill level of a person performing the task can be expressed by means of the execution capacity. It is clear that this measure of the skill level is the same for all values of the difficulty index. However, it is common that a person's skill level is quite high in easy tasks (ID is small), and low in difficult tasks; this suggests that some other criterion should be used to evaluate the skill level of a person. It should also be noticed that the execution capacity represents only one characteristic variable associated with the task that is used to determine the skill level.

In this report, we will study the evaluation of skill levels using the execution capacity of a person in a broader sense. Specifically, we will treat the execution capacity as a random variable

and specify it by means of the statistical average and variance. Then, we will introduce a new measure for skill evaluation, i.e., the probabilistic execution capacity of a person. It is then compared with that of a group of persons. In this way, we can evaluate the person's task performance relative to that of the average performances of the persons of the entire group.

We will then introduce a probabilistic skill index. It represents a novel way of evaluating the skills of a person performing a task. This approach can accomodate several variables which are used in the characterization of the person's skills.

2. DATA REPRESENTATION FOR A TASK

It is widely accepted that the magnitude (and variability) of the human responses and the information capacity of the human motor system play important roles in the evaluation of the human operators' motor skills. The human capacity to execute a task can be represented according to Fitts [1] by the index of task difficulty (ID) defined as

$$ID = \log_2 [A/(W/2)] \tag{1}$$

where A is the average amplitude of a human movement, and W is the admissible plus-and-minus ranges of terminal movement error ("target bandwidth").

It was conjectured in [1] that the average movement time (MT) should remain constant for different values of A and W (within limits) as long as the value of the ID remains constant. The ratio (ID)/(MT) expressed in bits per seconds may be interpreted as human's capacity for executing a particular class of motor responses. By postulating that motor movements follow a law similar to perceptual - motor processes, it is proposed in [1] that MT and ID are related as

$$(MT) = b + m \, (ID) \tag{2}$$

where m and b are constants. Equation (2) is often referred to as Fitts' law. In the framework of response duration, it may be considered as the traditional Weber function (the variability of a response as a function of its amplitude) written for the response duration. Subsequently performed experiments [1,2] have supported the relationship expressed in equation (2).

3. EXECUTION CAPACITY FOR SKILL EVALUATION

The constants m and b in equation (2) can be determined from experimental data. Indeed, if operator i, i = 1, 2,... consumes the movement time t_i in performing a task of difficulty index ID, and the same person has a reaction time of t_{oi} (i.e., when ID = 0, t_{oi} is the pure reaction time), then one may write the straight line equation that corresponds to equation (2) as

$$t_i = t_{oi} + m_i \ (ID) \tag{3}$$

The value of the reaction time t_{oi} is usually in the range 160 ms to 180 ms, which is small as compared with the movement time t_i.

It is a fairly common practice to use the inverse of the slope m_i, i.e., $1/m_i$ as a measure of the tracking skills of a person. If one considers t_{oi} negligible, then $1/m_i = (ID)/t_i$. If t_i is replaced by the average movement time, then this ratio may be interpreted as analogous to human's average capacity for executing a particular class of motor tasks. Its unit (bits/sec) is the same as that of the channel capacity in the information theory. Therefore, $1/m_i = C_i$ is referred to as the execution capacity of person i.

3.1 DETERMINATION OF EXECUTION CAPACITY

It is a usual practice to determine constants t_{oi} and m_i in equation (3) by fitting equation (3) to the (measured) data points using the least squared error method (regression analysis). It gives numerical values t_{oi}^{ls} and m_i^{ls} for the constants. Then, equation (3) specifies Fitts' law as a straight line for person i as

$$t_i^{ls} = t_{oi}^{ls} + m_i^{ls}(ID) \qquad (4)$$

The inverse $1/m_i^{ls}$ of the slope characterizes man's capacity C_i^{ls} to execute a particular task, i.e., human's execution capacity. A small value of $C_i^{ls} = 1/m_i^{ls}$ signifies that the movement time can become very large when the task difficulty index increases, i.e., the person has poor skills to execute the task. A large value of $C_i^{ls} = 1/m_i^{ls}$ corresponds to a small value of m_i^{ls} indicating that when a person is executing a task, the movement time does not vary much even though the difficulty of the task varies, i.e., the person has good skills for performing tasks even when the task difficulty index varies over a wide range. It is emphasized that only one variable ($C_i^{ls} = 1/m_i^{ls}$) is utilized here to measure human's capacity to execute a particular task.

Example 1

An exoskeleton comprised of seven rotary joints is attached to the arm of a human operator (Figure 1). It is moved by the human arm, and its joint motions are recorded by means of the encoders attached to the joint shafts. The exoskeleton arm is equipped with a force measuring pad attached to the elbow which can be used to compensate for the inertial torques. Moreover, a force/torque sensor (JR3) has been installed on the wrist of the exoskeleton to measure the forces/torques generated by the human hand. The joint shafts are connected via steel wires and

pulleys to the shafts of the motors. This system allows for the compensation of the gravity and the inertial torque effects of the exoskeleton. A peg of three different diameters may be attached to the gripper of the exoskeleton.

The task of a human operator to be considered here is to transfer the aforementioned peg from one hole to another specified hole as fast as possible. There are several holes in a vertical panel, and they all have the same diameter. Thus, the difference between the radius of a hole and that of the peg can be chosen by the selection of a particular peg size. One-half of this difference (tolerance) specifies W/2 in equation (1). The distance between the center points of the holes specifies A in equation (1). Hence, the task difficulty index can be calculated.

When a person is performing the described task, the time consumed for the completion of the task is measured electrically by means of the microswitches installed on the bottoms of the holes. A typical set of data points is shown in Table 1.

RUN#	ID	t_1
0	9.64	1711
1	8.06	671
2	9.06	1150
3	8.06	823
4	9.06	1602
5	9.64	1014
6	8.06	689
7	9.06	1465
8	8.06	770
9	9.06	1560
10	7.58	1028
11	6	733
12	7	768
13	6	678
14	7	1232
15	7.58	1085
16	6	619
17	7	1004
18	6	673
19	7	841
20	11.64	1397
21	10.06	692
22	11.06	1331
23	10.06	1373
24	11.06	1483
25	11.64	1626
26	10.06	904
27	11.06	1885
28	10.06	982
29	11.06	1798

Table 1

The data points of Table 1 for person $i = 1$ collected on the first day of the experiments are displayed in Figure 2 on the (t_1, ID) plane. A straight line is to be fitted to the data points using the regression analysis on equation (3). Thus, parameters t_{01}^{ls} and m_1^{ls} for equation (4) of person $i = 1$ and his execution capacity $C_1^{ls} = 1/m_1^{ls}$ for the task are to be calculated.

The least squared error method gives the numerical value of the parameter $m_1^{ls} = 151.7 \text{ms}/\text{bits}$. Thus, equation (4) assumes the following form:

$$y = 151.7(\text{ms}/\text{bits})x - 208.7(\text{ms}) \qquad (5)$$

where x = ID, the index of difficulty (in bits) and y = t_1 (in milliseconds), the movement time, which the person needs to complete the task.

Therefore, the execution capacity of person i = 1 is $C_i^{ls} = 1/m_1^{ls} = (1/151.7)\text{bits}/\text{ms} = 6.59(\text{bits}/\text{s})$ for the task in question. The calculated execution capacity is the same for all values of the task difficulty. It should be noted that the negative value of t_{o1}^{ls} is nonrealistic, although the regression analysis gives this value.

3.2 COMPARISON OF SKILLS USING EXECUTION CAPACITY

Fitts' law for the task described in Example 1 for person i is represented mathematically by equation (1). Person i repeated the experiments usually on six successive days. The samples measured on a particular day can be graphed on the $(t_i,(\text{ID}))$ plane. The regression analysis is applied to the data collected on that day to determine the constants in equation (1). It results in a set of linear equations for each day the experiment was performed; the equations are shown in the second column of Table 2. Then, the execution capacity C_i^{ls} for person i on the particular days is calculated.

As an example for person i = 1 the set of equations calculated are displayed in Table 2. Then, the execution capacity $C_i^{ls} = 1/m_{mi}^{ls}$ is calculated; they are displayed in the fifth column of Table 2 (i = 1).

Person $i =$	Equation (4) $y = t_i$, $x = ID$	Sample Average		Exec. Cap. C_i^{ls}	$\overline{C}_i^{ls} / \hat{\sigma}_{C_i^{ls}}$
		$\bar{t}_{oi} / \hat{\sigma}_{toi}$	$\hat{\bar{m}}_i / \hat{\sigma}_{mi}$		
1	y=151.7x-208.71 y=121.69x-26.124 y=131.37x-45.751 y=98.489x+37.134 y=105.33x+46.052 y=127.9x-80.906	-46.38/93.24	122.75/19.14	6.59 8.21 7.61 10.15 9.49 7.81	8.31/1.30
2	y=224.16x-275.17 y=173.75x+151.43 y=157.3x+11.88 y=315.62x-1086.2	-299.52/553.71	-217.71/71.20	4.46 5.75 6.35 3.16	4.93/142
3	y=157.06x-285.33 y=140.86x-221.88 y=66.838x+477.34 y=138.62x-146.73 y=134.23x-48.161 y=157.04x-289.41	-85.70/290.53	132.54/33.58	6.36 7.10 14.96 7.21 7.45 6.37	8.24/3.32
4	y=133.77x-70.574 y=137.47x-25.472 y=160.81x-204.18 y=146.46x-104.28 y=164.53x-227.78	-126.46/86.78	148.61/13.70	7.47 7.27 6.21 6.82 6.07	6.77/.62
5	y=248.46x-519.21 y=227.27x-357.11 y=312.27x-651.73 y=246.48x-360.75 y=165.74x-70.284 y=229.73x-330.5	-381.60/196.20	238.33/47.12	4.02 4.40 3.20 4.05 6.03 4.35	4.34/.93
6	y=93.86x+163.5 y=126.54x-79.094 y=81.912x+278.45 y=93.947x+231.33 y=102.59x+186.43 y=124.87x+2.27	130.48/139.03	103.95/18.10	10.65 7.90 12.20 10.64 9.74 8.00	9.86/1.67
7	y=89.686x+193.51 y=96.211x+203.51 y=114.44x-9.152 y=128.46x-57.104 y=87.679x+215.58	109.20/131.41	103.30/17.58	11.15 10.39 8.73 7.78 11.40	9.87/1.57

Table 2

The sample average of the execution capacity of person i = 1 is determined from the data: $\hat{\bar{C}}_1^{ls} = 8.31$ with the sample standard deviation $\hat{\sigma}_{c_1} = 1.30$ (in the sixth column of Table 2). If it is assumed that the range of the execution capacity for person i is specified as $(C_i^{ls} - \sigma_{c_i}, C_i^{ls} + \sigma_{c_i})$, then for person i = 1 an estimate for this range is (7.01, 9.61).

The comparison of the relative skills of the persons in the group can be performed by calculating $(\hat{\bar{C}}_i^{ls}, \hat{\sigma}_{c_i})$ for each person, and the corresponding values $(\hat{\bar{C}}^{ls}, \hat{\sigma}_c)$ for the entire group. An example is next presented to illustrate the approach.

Example 2

The execution capacity C_i^{ls} for person i = 1, .., 7 was calculated from the measured experimental data. The sample values of C_i^{ls} are displayed in the fifth column of Table 2. The sample average $\hat{\bar{C}}_i^{ls}$ was then calculated for each person i. The results are shown in the sixth column of Table 2.

The sample average of the execution capacity $\hat{\bar{C}}_i^{ls}$ of the entire population determined from the values of Table 2 is : $\hat{\bar{C}}^{ls} = 7.57$ and $\hat{\sigma}_c = 2.64$. Thus, the range of the execution capacity of the group is estimated as $(\hat{\bar{C}}^{ls} - \hat{\sigma}_c, \hat{\bar{C}}^{ls} + \hat{\sigma}_c) = (4.93, 10.21)$.

The skills of a particular person i can now be compared with the average skills of the group. For example, person i = 1 possess skills for the task under consideration which are higher than those of an average person in the group. Moreover, the range of the execution capacity for person i = 1 is (7.01, 9.61) which is entirely inside the range of the execution capacity of the entire group.

As another example, person i = 4 has skills for this task described by $\hat{\bar{C}}_4^{ls} = 6.77$ and $\hat{\sigma}_{c_4} = 0.62$. Thus, his skills in the average are below the average execution capacity $\hat{\bar{C}}^{ls} = 7.57$ of the group. His range of the execution capacity on the days of the experiments is (6.15, 7.39). A part of this range intersects with the range of the execution capacity of the entire group, while another part of this range is outside the range of the entire group. It can thus be expected that he has potential to perform on certain days on the average level of the other individuals in the group.

20-11

It should be noted that the described view applies over the entire range of the index difficulties of the task.

4. PROBABILISTIC CRITERION FOR SKILL EVALUATION

Equation (3) is a straight line on a plane with (t_i, ID) as the coordinates. The constants t_{oi} and m_i of the straight line are used to characterize the human performance. The values of these constants in equation (3) depend on the distribution of the measured sample values. Thus, m_i and t_{oi} may be considered as random variables. Indeed, it can be assumed that m_i and t_{oi} represent independent random processes.

4.1 PROBABILISTIC DESCRIPTION OF INDIVIDUAL SKILLS

If one performs the expectation operation on both sides of equation (3), the (statistical) mean $\bar{t}_i$ of the movement time t_i is

$$\bar{t}_i = \bar{t}_{oi} + \bar{m}_i(ID) \tag{6}$$

where $\bar{m}_i$ is the inverse of the mean task execution capacity for person i, i.e., $\bar{m}_i = E[m_i]$, and $E[\cdot]$ signifies the expectation operation. The average reaction time t_{oi} of person i is defined similarly. It is noted that equation (6) represents Fitts' law described also by equation (2).

Since it is assumed that t_{oi} and m_i are independent random variables, the variance σ_i^2 of the movement time t_i changes proportional to the square of the task difficulty (ID), i.e.,

$$\sigma_i^2 = \sigma_{toi}^2 + \sigma_{m_i}^2 (ID)^2 \tag{7}$$

where σ_{toi}^2 is the variance of t_{oi} and $\sigma_{m_i}^2$ the variance of m_i.

20-12

The values of the aforementioned means and variances can be estimated by calculating the sample means and the sample variances. If $\bar{\rho}_i$ signifies the expected value of t_{oi} or m_i, then an estimate for $\bar{\rho}_i$ when N sample values are given can be calculated by the following expression:

$$\hat{\bar{\rho}}_i(N) = \sum_{k=1}^{N} \bar{\rho}_i(k)/N \qquad (8)$$

where $\bar{\rho}_i(k)$ is the value determined in the k$\underline{th}$ experiment of person i, and $\hat{\bar{\rho}}_i(N)$ is the estimate of the statistical mean of $\bar{\rho}_i(N)$ based on the N sample values. Equation (8) can also be expressed in a recursive form as

$$\hat{\bar{\rho}}_i(N) = \hat{\bar{\rho}}_i(N-1) + [\bar{\rho}_i(N) - \hat{\bar{\rho}}_i(N-1)]/N \qquad (9)$$

The sample variance of ρ_i is calculated as

$$\hat{\sigma}_{\rho_i}^2(N) = \sum_{k=1}^{N} [\bar{\rho}_i(k) - \hat{\bar{\rho}}_i(N)]^2/(N-1) \qquad (10)$$

The sample means and sample variances (without denoting the argument N) will be used in the sequel for the average values of the random variables t_{oi} and m_i.

For a given task difficulty, i.e., when ID has a specified value, it may be assumed that the average movement time t_i of person i obeys a specific distribution with known (sample) averages $\bar{t}_{oi}^D$ and $\hat{\bar{m}}_i^D$, and (sample) standard deviations $\hat{\sigma}_{toi}^D$ and $\hat{\sigma}_{m_i}^D$, where the superscript "D" refers to a specific value of the difficulty level of the task. Such a probability density function can be displayed as shown in Figure 3. By equation (6), the mean value $\bar{t}_i$ of the movement time t_i for completing a given task increases linearly with the task difficulty ID, and by equation (7) the variance of t_i is proportional to the square of the task difficulty ID. Thus, the variance of t_i for

easy tasks (i.e. small values of ID) is relatively small. On the other hand, the values of t_i will be quite dispersed indicated by a large variance in difficult tasks (i.e., large values of ID).

4.2 PROBABILISTIC EVALUATION OF POPULATION SKILLS

For a given index of task difficulty, the statistical parameters of a probability density function for the movement time (t_i) used by an *individual* can be approximately determined on the basis of measured data. The corresponding parameters can also be calculated for a probability density function of the *group* of individuals. In order to compare the skills of a person with those of the entire group, the estimated probability density functions for persons i = 1, 2, ..., P and for the entire (available) population can be used as the basis of the evaluation.

The sample average movement time for the entire population of P persons can be estimated using available data points:

$$\hat{\bar{t}}^D = \sum_{i=1}^{P} \hat{\bar{t}}_i^D / P \tag{12}$$

where $\hat{\bar{t}}_i^D$ is the sample average movement time for individual i performing a task of specified difficulty index and $\hat{\bar{t}}^D$ is the corresponding time averaged for P persons. Similarly, the associated sample variance $(\hat{\sigma}^D)^2$ can be determined

$$(\hat{\sigma}_t^D)^2 = \sum_{i=1}^{P} [t_i^D - \hat{\bar{t}}^D]^2 / (P-1) \tag{13}$$

Thus, the sample mean $\hat{\bar{t}}^D$ and sample variance $(\hat{\sigma}_t^D)^2$ for the population of P persons can approximately be calculated. A density function for the process can then be estimated. It can then be utilized to evaluate the performance skills of the group and of a particular individual.

4.3 PROBABILISTIC SKILL INDEX FOR A SINGLE SKILL

It is will be assumed that an approximate density function for the "general" population has been determined. Then, a probabilistic measure for the execution skills of a group of persons performing a task of a given difficulty index can be defined:

Definition: The probabilistic skill index for a group of persons executing a task is the probability that the persons in the group can successfully perform the task of difficulty index $ID = d_1$ in time less than a_1.

The probability skill index defined specifies the performance of the group. Indeed, the skill index SI^D at the specified task difficulty level is mathematically expressed as

$$SI^D = Prob[t^D < a_1 | (ID) = d_1] \qquad (14)$$

where a_1 and d_1 are constants, and t^D is the movement time in the task of difficulty ID.

For example, in a "placing a peg in the hole" task, $a_1 = 899$ ms and $ID = 8$. Then $Pr[t^D < 899$ ms $| ID = 8]$ is the probability that the general population can perform a task of difficulty index 8 successfully in the (movement) time less than 899 ms. This probability can be evaluated since an approximate density function for the general population is known. Suppose that $Pr[t^D < 899$ ms $| ID = 8] = 0.70$.

The corresponding probabilistic skill index for individual i is similarly defined as

$$SI_i^D = Prob[t_i^D < a_1 | (ID) = d_1] \qquad (15)$$

The computations of SI_i^D gives the probabilistic skill index for a particular person. That is, $Prob[t_1^D < 899$ ms $| ID = 8]$ for an individual $i = 1$ can be calculated using his approximate density function with (sample) mean of $\hat{t}_1^D$ and (sample) standard deviation of $\hat{\sigma}_1^D$. Suppose that

20-15

Prob[t_1^D < 899 ms I ID = 8] = 0.80 for individual i = 1. Thus, the probabilistic skill index of this particular person is above that of an average population.

Suppose that the index of the task difficulty is ID = 9. Then the corresponding probabilities are smaller. For example, the probabilities for the successful completion of a task in less than 899 ms for the entire *group* and for the *individual* are Pr[t^D < 899 ms I ID = 9] = 0.45 and Prob[t_1^D < 899 ms I ID = 9] = 0.55, respectively.

It should be noticed that the probabilistic skill index introduced here will, in general, vary with the task difficulty index. Moreover, this skill index allows one to consider several variables in the evaluation.

4.4 PROBABILISTIC SKILL INDEX FOR A MULTIPLE SKILLS

Fitts' law describes the performance of an individual executing a task which is quantified by means of *one* variable. Due to its geometric nature, it is difficult to use in a task characterized by means of *several* variables. On the other hand, the probabilistic skill index introduced here can be used equally well to describe the performance of a person executing a task that is characterized using several variables. Indeed, a multivariate probability density function can be used in the multivariable case, and the basic approach described in Section 4.4 can be applied to calculate the probabilistic skill indices for multivariate cases as well.

The use of a probabilistic skill index can be demonstrated in the evaluation of a person's skills in a task that is characterized by two variables. For example, the execution of the task "peg into a hole" may be quantified by means of the time (t^D) as well as the energy (en^D) used by a human. Thus, the probabilistic skill index of individuals is then

$$SI^D = Prob[t_i^D < a_1, en^D < a_2 \text{ I } ID = d_1] \qquad (16)$$

The probability in equation (16) can be calculated using a multivariate probability density function. Indeed, using experimental data the mean and variance of the chosen multivariate density function can be estimated. Then, the probability skill index in equation (16) can be obtained. The data for such tasks are being currently collected in the Armstrong Laboratory; however, they were not available at the time of writing this report.

REFERENCES:

[1]   P. M. Fitts and J. R. Peterson, "Information Capacity of Discrete Motor Responses", *Journal of Experimental Psychology,* Vol 67, No. 2, February 1964, pp. 103-112.

[2]   P. M. Fitts, "The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement", *Journal of Experimental Psychology,* 1954, pp. 381-391.

Figure 1

rcv juldate 093 fcl

Capacity = 6.59

$y = 151.7x - 208.71$
$R^2 = 0.4912$
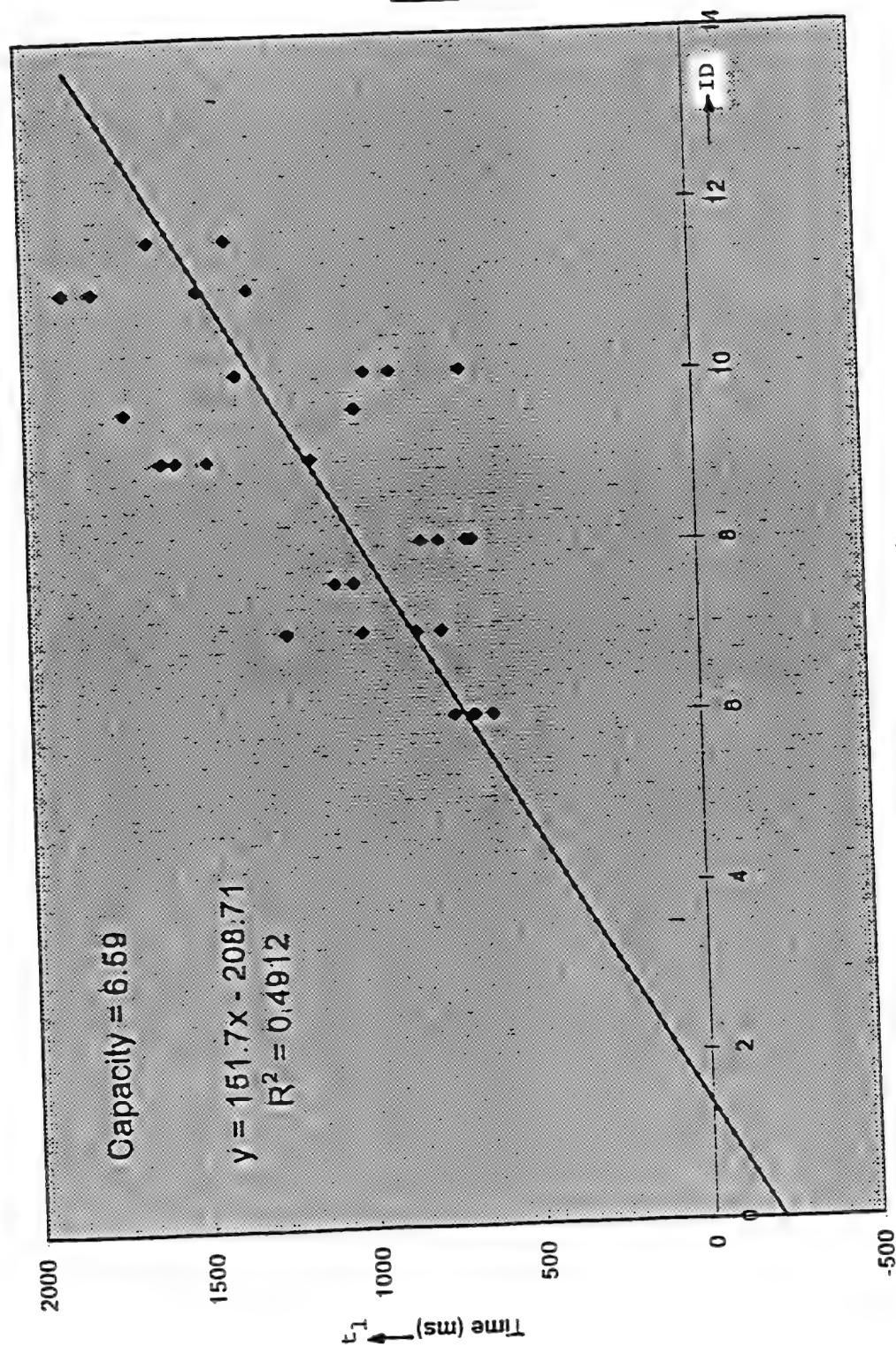
Bits of Difficulty

Figure 2
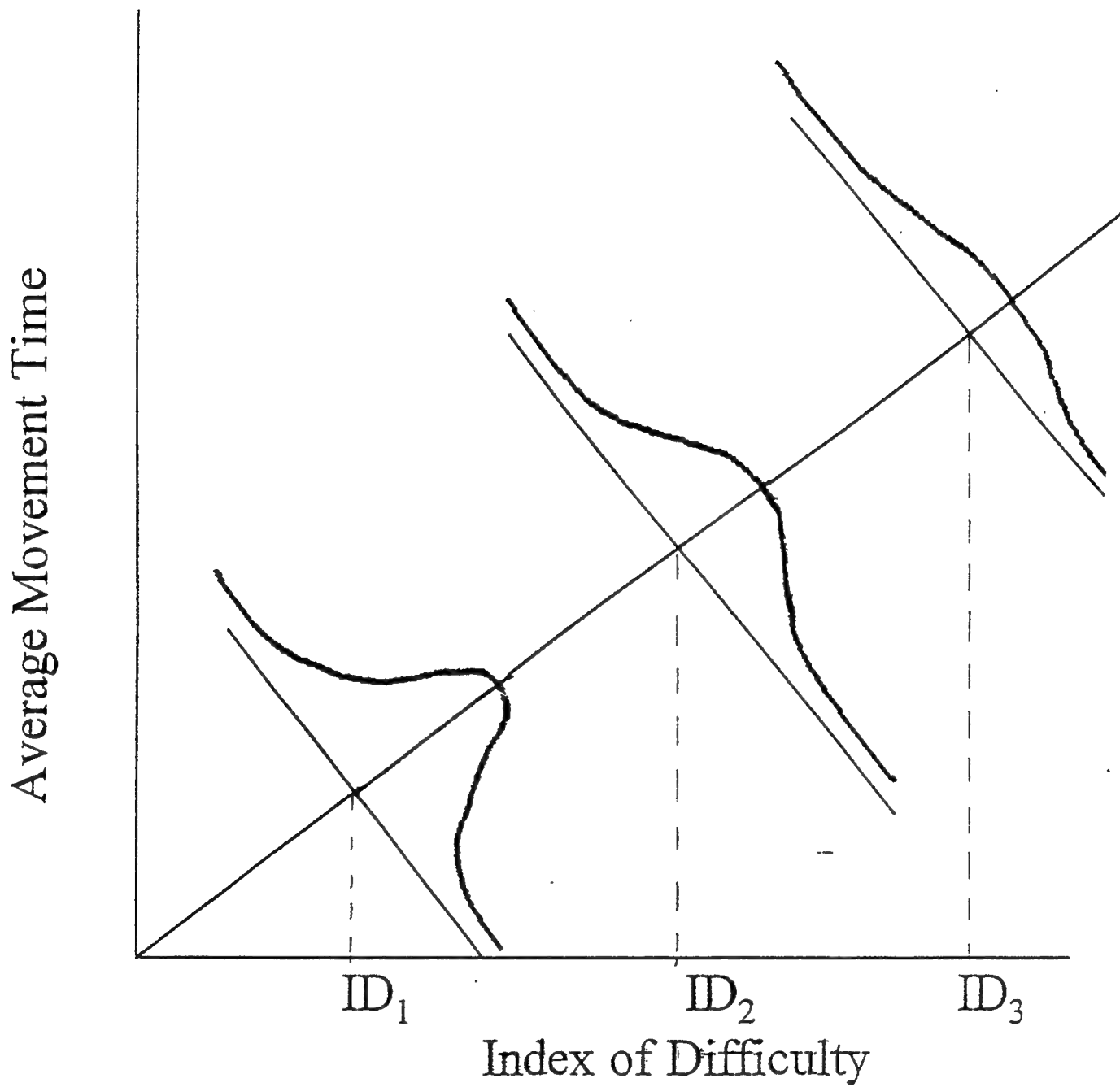
Figure 3

AROMATIC HYDROCARBON COMPONENTS IN DIESEL, JET-A AND JP-8 FUELS

S. Bin Kong, Ph.D.
Associate Professor
Department of Chemistry

University of the Incarnate Word
4301 Broadway
San Antonio, TX 78209

AROMATIC HYDROCARBON COMPONENTS IN DIESEL, JET-A AND J-8 FUELS

S. Bin Kong, Ph.D.
Associate Professor
Department of Chemistry
University of the Incarnate Word

Abstract

JP-8, Jet-A and diesel fuels were analyzed by Gas Chromatography-Masspectrometry(GC-MS) to identify the chemical components, especially aromatic hydrocarbons. These analyses were performed with GC cross-linked silicate capillary columns interfaced with a mass selective detector. JP-8 (Kelly A. F. B.) contains approximately 28 alkylated benzenes(11 %) and 13 naphthalenes(6%). Most of alkylbenzenes had molecular weights less than 150, such as ethylbenzene, dimethyl-, ethylmethyl- , methylpropyl-, ethyldimethyl-, trimethyl-, diethylbenzenes. Naphthalene derivatives were decalin, methyldecaline, methyl- dimethyl- and trimethylnaphthalenes. Jet-A (Berry Airfield) has 18 alkylated benzenes (8 %) and 12 naphthalenes(4%). Diesel fuel (Kelly A. F. B.) had 5% alkylated benzenes, 12 % naphthalenes and approximately 83 % hydrocarbons and others. We discovered 23 alkylated benzenes and 23 other aromatic hydrocarbon derivatives in diesel fuel. Mass spectra were effective to identify aromatic compounds but not accurate to determine positional isomers.

AROMATIC HYDROCARBON COMPONENTS IN DIESEL, JET-A AND JP-8 FUELS

S. Bin Kong, Ph.D.

Introduction

JP-8 and Diesel fuels from Kelly Air Force Base and Jet-A fuel from Berry Airfield were analyzed for alkylbenzene and naphthalene derivatives.

Aromatic hydrocarbons usually occur naturally and most of them are not harmful to human health. But some of them are known to be health hazards and are found throughout the environment in the air, water, and chemical waste site soil (Purcell, et al. 1990, Satcher, 1993). Like many other chemicals, we still are not aware the environmental or occupational health effects of many aromatic hydrocarbons at this time.

Petroleum contains many substituted benzenes, naphthalenes and other aromatic hydrocarbons and we are exposed to many of these substances by breathing, eating, drinking , or skin contact. It was reported that repeated contact with fuel oils can cause skin and liver cancer in mice. However, there is some conflicting information (Andrews and Snyder, 1991 and Henderson,1996).

Jet fuels and diesel fuels are complex mixtures of hydrocarbons produced by distillation of crude oil. The composition of fuels varies depending on the source of crude oil, distilling process, and specifications. The hydrocarbons in diesel and jet fuels are less volatile than those in gasoline. JP-8 is similar to commercial Jet-A fuel and consists primarily of C8 - C17 hydrocarbons with approximately 18 % aromatics. For diesel (marine), these are roughly 13 % paraffins, 43 % benzene aromatics, and 44 % naphthalenes (Henderson, 1996). It might also contain less than 10 % polycyclic aromatic hydrocarbons. Many compounds in the fuels do not exist in the vapors(Bishop, 1982).

Fuels can be released to the environment by in-flight exhaust and from spills and leaks to soil and water. Some of the aromatic hydrocarbons making up fuels may dissolve in water and then partition to sediment or be biodegraded. Some components may migrate through the soil in to the groundwater. It is not possible to fully describe the toxicokinetics of the individual hydrocarbons that are present in the petroleum (Satcher, 1993). However, it is important to identify the individual chemicals as a base for the future environmental studies.

Methodology

Diesel , JP-8 and additives were obtained from Kelly Air Force Base, San Antonio, Texas. Jet-A was obtained from Berry Airfield.

Analyses of the chemical components were done with GC crosslinked silicate capillary column interfaced with HP 5970 mass selective detector. Hydrocarbon standards were also used to identify n-paraffins and naphthalene-d8 was used as an internal standard. The injection mode was spitless and the column was temperature programmed (100 m x 0.25mm ID. Column temperature: 35 C for 15 minutes then to 230 C at 2 C /minute. He carrier gas flow rate was 50 mm/min.) (Supelco, 1996, Restek, 1996 and Brook, 1978).

Discussion

Electron Ionization (EI) was used for mass analysis. EI plays an important role in the routine analysis of small molecules. Its technique is straightforward and the databases contain the EI mass spectra of over 100,000 compounds and are used daily by thousands of chemists. However, it is difficult to analyze heavy molecules even if the backgrounds are subtracted due to the degree of fragmentation. It can limit the chemist to small molecules well below the mass range of common bioorganic compounds. Chemical Ionization(CI), electrospray ionization(ESI) or fast atom/ion bombardment(FAB) can be useful tools for future investigation. The usefulness of EI decreases significantly for compounds above a molecular weight of 400. The requirement

that the sample be thermally desorbed into the ionization source often leads to decomposition prior to vaporazation. The principal problems associated with thermal desorption in EI are the involatility of large molecules, thermal decomposition, and in many cases, excessive fragmentation. Too much fragmentation often results in no observable molecular ion. However, most of the aromatic hydrocarbons in fuels that we are dealing with are not huge molecules. Therefore, if we combine the data derived from EI and CI with GC/MS/IRD, that could be very useful for analyzing most of the aromatic hydrocarbons (Hafner et al., 1987 and Fodor, 1992).

Mass spectra

There were 182 peaks on mass spectra of JP-8 (Figure 1 and 2). JP-8 contains 11% benzenes and 6 % naphthalenes (Table 1). There were 28 benzenes and 13 naphthalene derivatives (Table 2). Most of the alkylbenzenes had molecular weight less than 150 such as ethylbenzene, dimethyl-, ethylmethyl-, methylpropyl-, ethyldimethyl-, trimethyl-, diethylbenzenes. Naphthalene derivatives were decalin, methyldecaline, methyl- dimethyl and trimethylnaphthalenes. It was difficult to identify the chemical components with mass spectra only in many cases due to the overlapping of the retention times of molecules. Even if we try to change the column conditions, carrier gas or flow rate, the interpretation can be tricky, despite making background subtractions. We should consider other ionization methods such as CI or methods for high molecular weight determination.

Jet-A has 8 % benzenes and 4 % naphthalenes. There were 18 alkylbenzene and 12 naphthalene derivatives in Jet-A (Table 3). Most of the alkylbenzenes were similar to that of JP-8 such as ethylbenzene, dimethyl-, ethylmethyl-, trimethyl-, ethyldimethyl-, methylpropyl-, isopropylmethyl- and trimethylbenzenes. Naphthalene derivatives were methyldecalin, methyldecahydronaphthalene, methyl-, ethenyl-, dimethyl and trimethylnaphthalenes.

Diesel fuel was identified with 5% benzenes, 12 % naphthalenes and approximately 83 % hydrocarbons and others. We discovered 23 benzenes and 23 other aromatic hydrocarbon derivatives in diesel fuel(Table 4). We found toluene, ortho- meta- para-xylenes, ethylmethyl-, phenylmethanal, trimethyl-, methylpropyl, ethyldimethyl-, 2-butyl, t-butyl- and trimethylindans. We also found isobutenyl-, dimethylpropenyl-, trimethylethenyl-, propenylbenzenes, but the characterization was difficult. We need high resolution MS data for the identification. Naphthalenes were dimethyl-, trimethyl-, dimethyltetrahydro-, trimethyltetrahydronaphthalenes. Azulene, biphenyl, 9-H-fluorene, and dimethylphenathrene were also found.

Conclusion

Identification of fuel components can be done more efficiently if we use GC/FT-IRD/MS. Chemical Ionization(CI) should be applied along with EI. To characterize biodegraded molecules which can have high molecular weights, FAB , electron spray(ESI) or other methods can also be considered. Mass spectra are effective to identify the aromatic hydrocarbons but are not accurate for determining positional isomers. To solve the problem above, we also need individual aromatic hydrocarbon standards in different concentrations.

Future Work

This summer, we have characterized fuels to identify the chemical components as a basis for future research. We will use these data for chemical and biodegradation studies in the following years. After we establish the routine procedure and identification of fuels, we can proceed to weathered JP-8 and other samples for degradation studies in chemical waste sites. Microorganism remediation study is also possible.

File   : C:\HPCHEM\1\DATA\F200696\JP8KELLY.D
Operator   :
Acquired   : 20 Jun 96   1:18 pm using AcqMethod FUELS
Instrument : 5970 - IN
Sample Name: jp-8 Kelly AFB
Misc Info :
Vial Number: 1

TIC: JP8KELLY.D
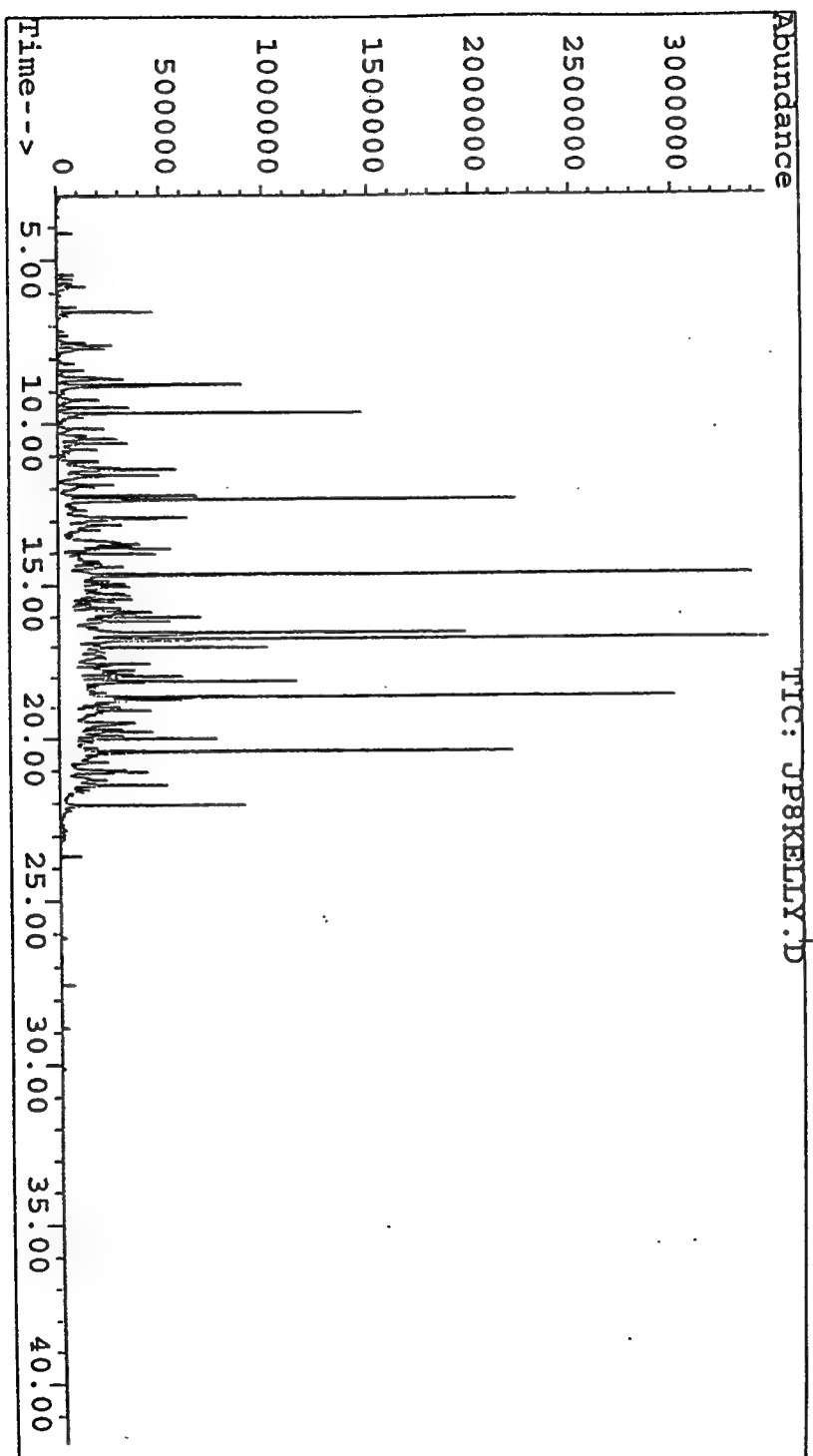
Figure 1. Gas Chromatogram of JP-8 fuel (Kelly A. F. B.)

Figure 2. Chemical Components of JP-8.

1	cyclohexane
2	?
3	methylcyclohexane
4	2-hexenol
5	pyrrolidine
6	toluene
7	?
8	cylohexane, 1,4-dimethyl-
9	?
10	?
11	?
12	?
13	cyclohexane, 1,2-dimethyl-
14	octane
15	cyclohexane, 1,4-dimethyl-
16	hexane, 2,2,5-trimethyl-
17	heptane, trimethyl-
18	?
19	?
20	heptane, 2,6-dimethyl-
21	cyclohexane, ethyl-
22	cyclohexane, 1,1,3-trimethyl-
23	nonene, 2,4,6-trimethyl-
24	?
25	cyclohexane, 1,2,4-trimethyl-
26	?
27	heptane, 2,4-dimethyl-1-
28	?
29	benzene, ethyl-
30	octane, 4-methyl-
31	octane, 1-methyl-
32	benzene, 1,4-dimethyl
33	?
34	?
35	? pentenal
36	?
37	? 1-hexene, 6-bromo
38	heptene, 2-
39	cyclohexane, 4-ethyl-methyl-
40	benzene, 1,3-dimethyl-
41	nonene

42	pentenal, 2-ethyl-
43	cyclohexane, 1-ethyl-3-methyl-
44	heptyne
45	decane
46	1-hexene, 3,5,5-trimethyl-
47	cyclohexanepropanaol
48	nonane, 4-methyl-
49	cyclohexane, propyl-
50	nonene, 3-methyl-
51	heptane, 3-ethyl-2-methyl-
52	3-nonen-1-ol
53	octyne
54	undecene
55	benzene, propyl-
56	pentane, tetramethyl-
57	nonene, 4-methyl-
58	benzene, 1-ethyl-2-methyl-
59	decanal
60	benzene, 1,2,4-trimethyl-
61	heptane, propyl-
62	benzene, 1-ethyl-2-methyl-
63	cyclohexane, 1-ethyl-2-methyl-
64	4-cyclohexyl-1-butanol
65	pentanal, ethyl-
66	benzene, 1,2,4-trimethyl-
67	decane
68	benzene, butyl-
69	?
70	benzene, methylpropyl-
71	decane. 2,.5,6-trimethyl-
72	decane, 4-methyl-
73	benzene, 1-ethyl-4-methyl-
74	decane, dimethyl-
75	cyclohexane, 1-methylpropyl-
76	?
77	cyclohexane, butyl-
78	4-pentenal, 2-methyl-
79	6-octenal, 3,7-dimethyl-
80	2-decenal
81	benzene, 2-ethyl-1,3-dimethyl
82	benzene, 1-methyl-3-propyl

83	naphthalene, decahydro-, trans-	
84	decane, 5-methyl-	
85	benzene, 1-ethyl-2,3-dimethyl	
86	decane, 2-methyl	
87	decane, 3-methyl	
88	nonenal	
89	nonadiene, dimethyl	
90	? benzene, dimethylethyl	
91	benzene, 1-ethyl-2,3-dimethyl order	not in
92	, 4-ethyl-2,3-dimethyl , 2-ethyl-1,3-dimethyl cyclohexane, 1-ethyl-1-methyl-	
93	cyclohexane, 1-methyl-1-propyl benzene, 2-ethyl-1,3-dimethyl	
94	benzene, 1-ethyl-2,3-dimethyl cyclohexane, 1-ethyl-2-propyl	
95	benzene, 4-ethyl-1,2-dimethyl	
96	undecene	
97	benzene, [1,1-dimethylpropyl]-	
98	?	
99	octane, 2-methyl	
100	benzene, 1,2-diethyl	
101	benzene, 1-methyl-2-(1-methylethyl)-	
102	dodecane, 2,6,10-trimethyl	
103	naphthalene, decahydro-2-methyl	
104	? dodecanol, trimethyl	
105	cyclohexane, pentyl	—
106	benzene, 4-ethyl-1,2-dimethyl	
107	? azulene	
108	? benzene, 1,4-dimethyl-2-[2-methylpropyl]-	
109	? 2-hexyl-1-decanol	
110	? undecene, 5-methyl	
111	decane, 2,3-dimethyl	
112	? heptadecane, methyl	
113	? dodecane, 3-methyl	
114	benzene, [1,1-dimethylprophy]-	
115	? tetradecyne	
116	naphthalene d8 standard	
117	azulene	

118	1H-indene, 1-methylene- naphthalene dodecane
119	benzene, e-ethenyl-1,3,5-trimethyl
120	?
121	undecane, 2,6-dimethyl
122	? decane, 2,2,3-trimethyl
123	1H-indene, 2,3-dihydro-1,1,5-trimethyl-
124	decane, 2,3,7-trimethyl
125	? phosphoric acid, dioctadecyl ester
126	naphthalene, decahydro-2.6-dimethyl-
127	?
128	cyclohexyl, hexyl
129	1-tridecanol
130	? octadecyne
131	? decane, 2,6,7-trimethyl
132	dodecane, 5-methyl
133	? heptane, 2,4-dimethyl
134	nonadecane
135	dodecane, 3-methyl
136	nonane, 3-methyl
137	? octanol, hexyl
138	? octadecane, 1-chloro
139	?
140	? tridecanediol, diacetate
141	1-hexadecanol
142	tridecane
143	naphthalene, 1-methyl
144	? 6-octene-1-ol, 3,7-dimethyl-
145	?
146	benzocycloheptatriene
147	benzene, [3-methylcyclopentyl]
148	? pyridinol, acetate(ester)
149	?
150	? 7-hexadecyne
151	? tridecane, methyl
152	?
153	eicosane, 10 methyl
154	pentacosane
155	10-methylnonadecane
156	dodecane, 2,6,10-trimethyl

157 ?
158 naphthalene, 2-ethenyl
159 biphenyl
     tetradecane
160 octadecane, 1-chloro
161 naphthalene, 1,2-dimethyl
162 ?
163 octadecanol
164 naphthalene, 2,7-dimethyl
165 naphthalene, 2,6-dimethyl
166 ?
167 ? 2-hexyl-1-decanol
168 octylcyclohexane
169 ? quinolinone, hydrazone
170 pentadecane
171   butane, 2-iodo-2-methyl
172 3-dodecyne
173 ? decane, 3-methyl
174 ?
175 ? 1,2'-biphenyl-2-methyl
176 ? methylphenylhexatriene
177 ?
178 tetradecane
179 ?
180 ?
181 ?
182 ?

Table 1. Alkylbenzenes and naphthalenes of JP-8 (Kelly A. F. B.),
Jet-A (Berry Airfield), and Diesel (Kelly A. F. B.).

Benzenes & Naphthalenes of Fuels			
	JP-8	Jet-A	Diesel
Benzenes	10.63	8.25	5.34
Naphthalenes	6.35	4.17	11.80
Hydrocarbons & others	83.02	87.58	82.86
Total	100.00	100.00	100.00
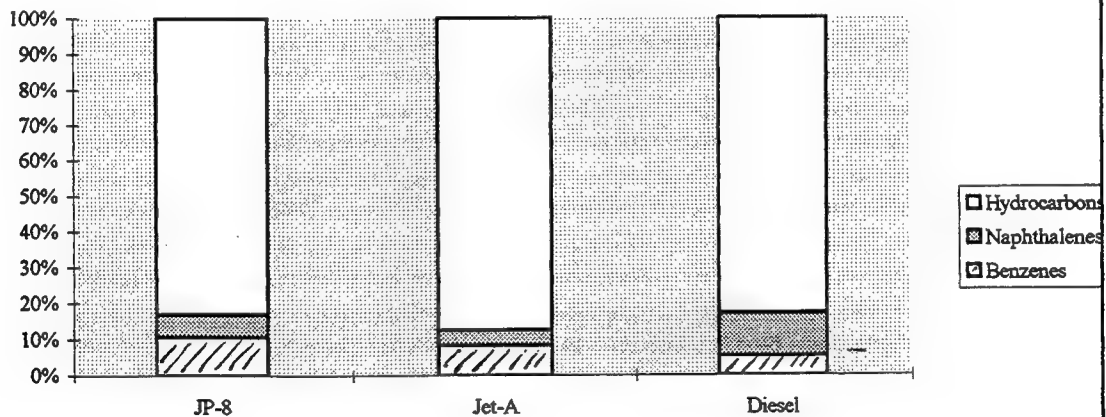
Table 2. Alkylbenzenes and naphthalenes of JP-8.

Benzenes and Naphthalenes in JP-8		
12 5.58 0.16 C:\DATABASE\NBS75K.L	CAS #	area%
Toluene	000108-88-3	0.16
Ethylbenzene	000100-41-4	0.31
Benzene, 1,3-dimethyl-	000108-38-3	2.73
p-Xylene	000106-42-3	
Benzene, 1,2-dimethyl-	000095-47-6	0.68
Benzene, 1-ethyl-2-methyl-	000611-14-3	1.44
Benzene, 1-ethyl-3-methyl-	000620-14-4	
1,2,4-Trimethylbenzene	000095-36-3	1.3
Benzene, 1,2,3-trimethyl-	000526-73-8	1.58
Benzene, butyl-	000104-51-8	0.14
Benzene, 1-ethyl-4-methyl-	000622-96-8	0.59
Benzene, 1-methyl-2-(1-methylethyl)-	000527-84-4	0.37
Benzene, 1-methyl-3-(1-methylethyl	000535-77-3	
Benzene, methyl(1-methylethyl)-	025155-15-1	
Benzene, 2-ethyl-1,3-dimethyl-	002870-04-4	0.21
Benzene, 1-methyl-4-(1-methylethyl	000099-87-6	
Benzene, 4-ethyl-1,2-dimethyl-	000934-80-5	0.78
Benzene, 1-ethyl-2,3-dimethyl-	000933-98-2	
Benzene, (1,1-dimethylpropyl)-	002049-95-8	0.2
Benzene, 1-ethyl-2,4-dimethyl-	000874-41-9	
Benzene, 1,2-diethyl-	000135-01-3	?
Benzene, 1,2,3,4-tetramethyl-	000488-23-3	?
Benzene, 1,2,4,5-tetramethyl-	000095-93-2	?
Naphthalene, decahydro-2-methyl-	002958-76-1	1.23
Benzene, 1,3-diethyl-5-methyl-	002050-24-0	0.63
Benzene, 1-ethyl-3,5-dimethyl-	000934-74-7	
Benzene, (1,1-dimethylpropyl)-	002049-95-8	0.31
Naphthalene-d8-	001146-65-2	
1H-Indene, 2,3-dihydro-1,1,5-trime	040650-41-7	? 0.28
1H-Indene, 2,3-dihydro-1,1,3-trime	002613-76-5	?
Benzene, 1-(1-methylethenyl)-3-(1-methyl	001129-29-9	?
Naphthalene, 1-methyl-	000090-12-0	1.15
Benzocycloheptatriene	000264-09-5	?
Naphthalene, 2-methyl-	000091-57-6	0.93
1H-Indene, 2,3-dihydro-1,1,3-trime	002613-76-5	
Naphthalene, 1,2,3,4-tetrahydro-1-methyl-	004175-54-6	0.09

21-14

Name	CAS#	%
1H-Indene, 2,3-dihydro-1,4,7-trime	054340-87-3	
Naphthalene, 2-ethenyl-	000827-54-3	0.32
Biphenyl	000092-52-4	
Naphthalene, 1,2-dimethyl-	000573-98-8	0.1
Naphthalene, 2,3-dimethyl-	000581-40-8	
4-Hexenoic acid, 3-methyl-2,6-diox	056771-77-8	?
Naphthalene, 2,6-dimethyl-	000581-42-0	0.73
Naphthalene, 1,4-dimethyl-	000571-58-4	0.55
Naphthalene, 2,7-dimethyl-	000582-16-1	0.87
Naphthalene, 1,3-dimethyl-	000575-41-7	
Naphthalene, 1,4,6-trimethyl-	002131-42-2	0.07
Naphthalene, 2,3,6-trimethyl-	000829-26-5	

Table 3. Alkylbenzenes and napthalenes of Jet-A.

Benzenes and Naphthalenes		
Jet-A Fuel, Berry		
NAME	CAS#	%
Ethylbenzene	000100-41-4	0.15
Benzene, 1,2-dimethyl-	000095-47-6	0.64
p-Xylene	000106-42-3	
Benzene, 1-ethyl-3-methyl-	000620-14-4	1.16
Benzene, 1-ethyl-4-methyl-	000622-96-8	0.68
Benzene, 1-ethyl-2-methyl-	000611-14-3	0.40
1,2,4-Trimethylbenzene	000095-36-3	1.32
Benzene, 1,2,3-trimethyl-	000526-73-8	0.25
Benzene, 1-methyl-3-propyl-	001074-43-7	0.83
Benzene, (1-methylpropyl)-	000135-98-8	0.10

Benzene, 4-ethyl-1,2-dimethyl-	000934-80-5	0.28
Benzene, 1-ethyl-2,4-dimethyl-	000874-41-9	0.18
Benzene, 2-ethyl-1,3-dimethyl-	002870-04-4	0.20
Benzene, 1-ethyl-2,3-dimethyl-	000933-98-2	0.97
Benzene, 1-methyl-3-(1-methylethyl	000535-77-3	0.69
Benzene, 1-methyl-2-(1-methylethyl	000527-84-4	0.40
Naphthalene, decahydro-2-methyl	002958-76-1	0.47
Naphthalene, 1,2,3,4-tetrahydro-1-methyl-	004175-54-6	0.21
Octadecanal	000638-66-4	
Naphthalene, 2-methyl-	000091-57-6	0.36
Naphthalene, 1-methyl-	000090-12-0	0.20
Naphthalene, 2-ethenyl-	000827-54-3	0.44
Naphthalene, 2,7-dimethyl-	000582-16-1	0.48
Naphthalene, 2,6-dimethyl-	000581-42-0	0.72
Naphthalene, 1,3-dimethyl-	000575-41-7	0.48
Naphthalene, 1,6-dimethyl-	000575-43-9	0.1
Naphthalene, 1,5-dimethyl-	000571-61-9	0.36
Naphthalene, 1,6,7-trimethyl-	002245-38-7	0.35
Naphthalene, 2,3,6-trimethyl-	000829-26-5	
Naphthalene, 1,4,6-trimethyl-	002131-42-2	

Table 4. Alkylbenzenes and naphthalenes of Diesel.

Benzenes and Naphthalenes		
Diesel Fuel , Kelly A. F. B.		
NAME	CAS#	%
Ethylbenzene	000100-41-4	0.11
p-xylene	000106-42-3	0.30
Benzene, 1,2-dimethyl-	000095-47-6	0.03
Benzene, propyl-	000103-65-1	0.01
Benzene, 1-ethyl-4-methyl-	000622-96-8	0.42
Benzene, 1-ethyl-2-methyl-	000611-14-3	0.34
Benzene, 1-ethyl-3-methyl-	000620-14-4	0.39
1,2,4-Trimethylbenzene	000095-36-3	0.47
Benzene, 1,2,3-trimethyl-	000526-73-8	
Benzene, 1-methyl-3-propyl-	001074-43-7	0.52
Benzene, 1-ethyl-2,3-dimethyl-	000933-98-2	0.49
Benzene, 1-ethyl-2,4-dimethyl-	000874-41-9	0.2
Benzene, 4-ethyl-1,2-dimethyl-	000934-80-5	0.12
Benzene, 1-methyl-3-(1-methylethyl	000535-77-3	0.11
Benzene, 1-ethyl-2,3-dimethyl-	000933-98-2	0.78
Benzene, 2-ethyl-1,3-dimethyl-	002870-04-4	
Benzene, 4-ethyl-1,2-dimethyl-	000934-80-5	
Benzene, 1-methyl-2-(1-methylethyl	000527-84-4	0.19
Benzene, tert-butyl-	000098-06-6	0.08
Naphthalene, 1,2,3,4-tetrahydro-	000119-64-2	1.08
Naphthalene-d8- standard	001146-65-2	7.17
Naphthalene, 1,2,3,4-tetrahydro-2-methyl-	003877-19-8	2.44
Naphthalene, 1,2,3,4-tetrahydro-1-methyl-	001559-81-5	0.5
Naphthalene, 1,2,3,4-tetrahydro-6-methyl-	001680-51-9	1.21
Naphthalene, 1,2,3,4-tetrahydro-5-methyl	001680-51-9	
Naphthalene, 1-methyl-	000090-12-0	1.15
Naphthalene, 1,2,3,4-tetrahydro-2,7-dimethyl-	013065-07-1	1.71
Benzene, 1-(1-methylethenyl)-2-(1-methylethyl)-	005557-93-7	0.24
Biphenyl		0.54
Naphthalene, 2-ethenyl-		
Naphthalene, 1,2,3,4-tetrahydro-1,5-dimethyl-	021564-91-0	1.31
Naphthalene, 2,6-dimethyl-	000581-42-0	0.35
Naphthalene, 2,7-dimethyl-	000582-16-1	0.87
Naphthalene, 1,6-dimethyl-	000575-43-9	1.08

Naphthalene, 1,2,3,4-tetrahydro-1,1,6-trimethyl-	000475-03-6	0.05
Naphthalene, 1,2,3,4-tetrahydro-1,5,7-trimethyl-	021693-55-0	0.05
Indan,1,1,6,7-tetramethyl-	016204-58-3	
Naphthalene, 2,3,6-trimethyl-	000829-26-5	0.45
Naphthalene, 1,4,6-trimethyl-	002131-42-2	
Naphthalene, 1,6,7-trimethyl-	002245-38-7	
Naphthalene, 1,4,5-trimethyl-	002131-41-1	
9-H-Fluorene, 2,3-dimethyl		0.01
Phenanthrene, 2.5-dimethyl-	003674-66-6	0.30

References

Andrews, L. S., and Snyder, R., 1991. Toxic effects of solvents and vapors. p681-722 in Casarett and Doull's Toxicology: The basic Science of Poisons, 3rd Ed., Klassen, C. D., Amdur, M. O., and Doull, J., eds. New York: Pergamon.

Bishop, E. C., 1982. Evaluating health hazards associated with aircraft fuel cell maintenanc. Proceedings of the twelveth conference on environmental toxicology, AMRL-TR-79-68. Wright-Patterson Air force Base, Dayton, Ohio.

Brook, J. J., 1978. Jet Engine exhaust analysis by subtractive chromatography, USAF Report SAM-TR-37.

Fodor, G. E., and Naegeli, 1992. Development of a method to determine the autoxidation of turbine fuels, AD A260578, Southwest Research Institute, San Antonio, Texas.

Hafner, A. M., Norton, K. L., and Griffiths, P. R., 1987. Interfaced gas chromatography and fourier transform infrared transmission spectrometry by elute trapping at 77K, Analytical Chemistry, 60, 2441.

Henderson, R. F., 1996. Permissible exposure levels for selected military fuel vapors, p18, National Academy Press, Washington, D. C.

Purcell, K. J., Cason, G. H., Gargas, M. L., and Anderson, M. E., 1990. In vivo metabolic interactions of benzene and toluene, Toxicol. Lett. 52, 141.

Restek Corporation, 1996. Product Guide, Bellefonte, PA.

Satcher, D., 1993. Toxicological profile for fuel oils, U.S. Dept. of Health & Human Services. Contract No. 205-93-0606.

Supelco, 1996. Chromatography Products, Supelco Park, Bellefonte, PA.

USAF OEHL Report 84-148SZ111CFC, 1984. Characterization and evaluation of JP-4, Jet A and mixtures of these fuels in environmental water samples, Suffolk County Airport, N. Y..

Acknowledgement

MENTAL WORKLOAD CLASSIFICATION VIA PHYSIOLOGICAL SIGNAL PROCESSING: EOG AND EEG ANALYSES

Xuan Kong, Ph.D.
Assistant Professor
Department of Electrical Engineering

Northern Illinois University
DeKalb, IL 60115

MENTAL WORKLOAD CLASSIFICATION VIA PHYSIOLOGICAL SIGNAL PROCESSING: EOG AND EEG ANALYSES

Xuan Kong, Ph.D.
Assistant Professor
Department of Electrical Engineering
Northern Illinois University

ABSTRACT

Accurate and efficient mental workload classification and assessment based on physiological data have many important applications ranging from operator monitoring to interaction and control of man/machine systems. The physiological signals studied in this project are the electrooculogram (EOG) and electroencephalogram (EEG). An efficient and robust eyeblink detection algorithm was developed based on nonlinear analysis of the EOG signal. The effectiveness of the algorithm was demonstrated via the analysis results of several benchmark test data sets provided by the Armstrong Laboratory. To quantify EEG changes associated with different mental workload conditions, a parametric model (autoregressive model) was established for each properly segmented EEG. The spatial relations among 57 channels of EEG were examined using Itakura distance. Preliminary analysis results indicate that the spatial distance maps exhibit unique patterns for various mental workload conditions. A workload indicator is then developed to capture the distinct distance patterns and data analysis results show promising classification results.

MENTAL WORKLOAD CLASSIFICATION VIA PHYSIOLOGICAL SIGNAL PROCESSING: EOG AND EEG ANALYSES

Xuan Kong, Ph.D.

This report contains two parts: EOG analysis and EEG analysis. Electrooculogram (EOG) and electroencephalogram (EEG) represent two frequently used physiological signals for mental workload assessment. The first part of this report summarizes a new EOG analysis algorithm developed for automatic detection of eyeblinks. The second part discusses a multi-channel EEG analysis using parametric model and Itakura distance measure.

PART I: EOG Study

1.1 Introduction

Eyeblinks are periodic, symmetrical closures and reopenings of the eyelid. To discriminate an eyeblink and an eye closure followed by a reopening of the eyes, durations of the eye closure are often used as an important criterion. A typical eyeblink closure lasts for less than 300 ms. Electrooculogram (EOG) records changes in the difference of electrical potentials between the cornea and retina as the eye movement occurs [3]. Therefore the EOG can be used to detect eyeblinks. Here, a new nonlinear algorithm is proposed to automatically detect eyeblinks based on EOG measurement.

In traditional physiological signal processing, EOG is often considered as noise or interference. Thus a great deal of efforts have been devoted to eliminate the EOG (or its effects). For example, eye movements (not necessarily eyeblinks) cause EOG signal to leak into the electrodes intended for electroencephlagram (EEG) recording. Thus the EEG electrodes are actually recording EEG+EOG signal. When both EOG and EEG+EOG signals are available, signal processing techniques can be developed to remove EOG from EEG+EOG in order to get a clean copy of the EEG.

There is a growing interest in exploring the information contained in eye movement patterns. For example, it is found that when a person is engaged in activities requiring great concentration (such as landing an airplane), eyeblink rate drops significantly [2]. To monitor or assess such a mental workload condition, it is desirable to develop automatic eyeblink detection algorithms. Since EOG can be conveniently recorded, it becomes a suitable candidate based on which eyeblink events can be detected. However, this is not a trivial task as eyeblink pattern varies significantly: Figure 1 shows EOG traces of four different eyeblinks and such variability in EOG justifies the need for developing EOG-based eyeblink detection algorithms which are robust to noise, inter-subject and intra-subject variations.

1.2 Pre-filtering of EOG

The single most important characteristics of the EOG for an eyeblink is the steep negative slope corresponding to the rapid eye closure. In order for an eyeblink algorithm to reliably detect such an event, a nonlinear filter called erosion is first applied to the EOG signal $x[k]$ (see Figure 2):

$$y[k] = \min_{e \in [-E,E]} x[k-e]. \tag{1}$$

where E is the order of the erosion filter. Unlike many linear filters, erosion filter will not smooth out the edges in the original signal (especially the leading negative slope). In fact, as one can see from Figure 2, the valley of the eyeblink waveform complex is actually enhanced. Another advantage of the erosion filter is that it can filter out positive spike-like noises.
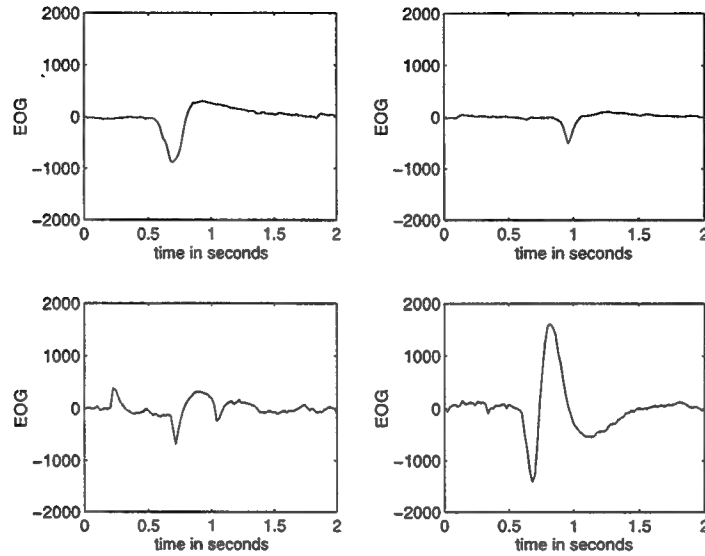
Figure 1: Sample EOG traces for four eyeblinks from the same subject. The figure demonstrates the large variations of EOG patterns corresponding to eyeblinks and the need for robust eyeblink detection algorithms.
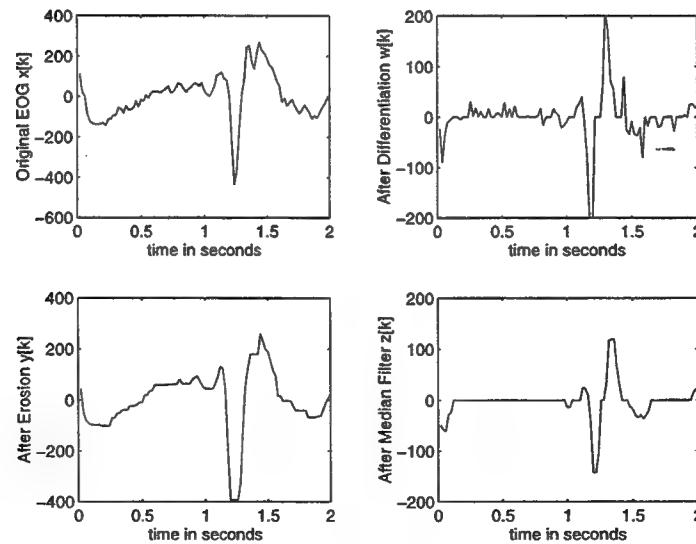

Figure 2: Intermediate EOG processing results of the new eyeblink detection algorithm.

Following the erosion, the slopes of possible eyeblink EOG waveform complex are determined. An eyeblink is characterized as a large negative slope followed by a positive slope of comparable size within a pre-defined time limit. A differentiation filter is employed to obtain the slopes of the EOG waveform. A "low-noise" Lanczos differentiation filter was introduced in [4]. The coefficients of a D-th order Lanczos differentiation filter are as follows:

$$h_k = -\frac{3k}{N(N+1)(2N+1)}, \quad k = -N, -(N-1), \cdots, N-1, N. \tag{2}$$

The first order differentiation has the familiar form of symmetric difference

$$w[k] = \frac{1}{2}(y[k+1] - y[k-1]). \tag{3}$$

To reduce the noise effects, now characterized by small fluctuations around zero, a median filter is used to average out these fluctuations and the order of the median filter is denoted as M (An M-th order median filter sorts $2M+1$ samples). To reduce the false alarm rate, a threshold filter is used to zero-out those slope values close to zero:

$$z[k] = \begin{cases} w[k] & |w[k]| \geq W \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

1.3 Eyeblink Detection Algorithm

Now detecting eyeblinks from EOG becomes determining in $z[k]$ a qualified sequence of a negative peak followed by a positive peak within a pre-defined time window. We note that the areas under the negative and positive peaks in $z[k]$ correspond to the magnitude of the negative and positive slopes in the original EOG signal and the peaks roughly correspond to the midpoints of eye closure and reopening which define the eyeblink duration [1].

Pseudo-codes of the above algorithm are shown in Figure 3. The algorithm was implemented in MATLAB and tested on a PC system. The MATLAB codes can be obtained from Dr. Glenn F. Wilson (USAF Armstrong Laboratory, AL/CFHP, 2255 H Street, Wright-Patterson AFB, OH 45433-7022). The MATLAB implementation was designed to analyze data conforming the WAM (Workload Assessment Monitoring) system format and can be adapted easily to process any other data format. The EOG data available was sampled at 500 Hz and such a high time resolution is not necessary for eyeblink detection. The EOG data was sub-sampled ten times so that the EOG analysis algorithm effectively processes data with a sampling rate of 50 Hz.

There are two modes under which the algorithm can operate: learning mode and detection mode. In practice, a flag can be used to allow the algorithm to switch automatically from the learning mode to the detection mode after the algorithm parameters are learned and adapted to the eyeblink patterns of a given subject.

In the learning mode the algorithm uses the default parameter settings to detect eyeblinks. The algorithm then does a statistical analysis of the eyeblink durations and sizes. The algorithm then adjusts the detection parameters accordingly.

In the detection mode, the algorithm first filters the EOG signal via erosion to enhance the features for eyeblink detection. The signal is then subject to differentiation (and thresholding) to minimize any baseline shifts. The algorithm then detects the sequence of a negative slope followed by a positive slope. The sizes of the slopes are then checked against the proper threshold values, as well as the durations of the positive and negative slopes and the time difference between them (corresponding to the blink duration). One important feature of this algorithm is that it uses a soft-decision process to calculate the likelihood of an event being a blink or non-blink based on each feature, instead of just accepting or rejecting the event outright. A composite score is reached once all the features are considered. The algorithm then determines whether the sequence of the events is a blink or not.

Another advantage of the composite score is that one can classify events as blinks, small blinks, possible blinks, or non-blinks, much like an expert would do when he or she inspects the EOG traces. Such a

```
Algorithm Initialization
WHILE more data to process DO
    Pre-processing:
        erosion filtering
        differentiation filtering
        median filtering
    IF a negative slope followed by a positive slope
        continue to process the data
    ELSE
        process next data point
    ENDIF
    Feature Evaluation:
        blink duration
        negative slope size
        positive slope size
        inter-blink distance
    IF features satisfy blink criteria
        mark data as a small or regular blink
    ELSE
        process next data point
    ENDIF
ENDWHILE
```

Figure 3: Pseudo-code for the new eyeblink detection algorithm.

classification scheme also provides a richer set of information for the decision process which follows. For example, eyeblink rate is an important factor in mental workload assessment. The assessment accuracy may be improved by taking into consideration of the more precise description of the eye movement like small blinks and regular blinks.

1.4   Algorithm Complexity

The computational complexity of the eyeblink detection algorithm is given as follows (per sample point):

operation	number
multiplication	15
addition	14
logic comparison	$11 + 2E + (2M + 1)\log_2(2M + 1)$

With the speed of 486 or 586 PCs, the algorithm can easily process the EOG for blink detection using only a small fraction of the computing resources. The current MATLAB implementation on a 486/100MHz PC processes the EOG data four times faster than real time, even with extensive graphic displays.

If the above computational complexity cannot satisfy the real-time operation requirement for a given hardware specification, pre-filtering portion can be incorporated into the (simpler) eyeblink detection algorithm of the WAM system. This augmented algorithm should have a better detection accuracy while the computational complexity of the algorithm is increased by only (per sample point):

operation	number
addition	1
logic comparison	$1 + 2E + (2M + 1)\log_2(2M + 1)$

EOG	Total	WAM System				New Algorithm		
Segment	Blinks	Miss	False	Training Data	Miss	False	Training Data
grc	33	6%	6%	100 seconds	6%	6%	100 seconds
grc1	34	3%	3%	120 seconds	6%	0%	100 seconds
grc2	14	7%	14%	whole segment (128 s)	14%	7%	100 seconds
grc3	14	0%	0%	whole segment (240 s)	7%	7%	100 seconds
the1	93	9%	17%	200 seconds	14%	9%	100 seconds
the2	148	9%	26%	200 seconds	9%	11%	100 seconds

Table 1: Performance comparison of the WAM system and the new eyeblink detection algorithm with respect to the standard of expert I.

EOG	Total	WAM System				New Algorithm		
Segment	Blinks	Miss	False	Training Data	Miss	False	Training Data
grc	29	3%	17%	100 seconds	0%	10%	100 seconds
grc1	31	6%	10%	grc training result	3%	6%	grc training result
grc2	9	0%	66%	grc training result	0%	44%	grc training result
grc3	10	0%	40%	grc training result	0%	10%	grc training result
the1	85	4%	25%	100 seconds	7%	16%	100 seconds
the2	131	7%	44%	120 seconds*	5%	18%	the1 training result

Table 2: Performance comparison of the WAM system and the new eyeblink detection algorithm with respect to the standard of expert II. *The WAM system eyeblink detection error becomes extremely high when using the training result for "the1," so the WAM system is trained again based on the new segment "the2."

1.5 Data Analysis Results

To evaluate the performance of the eyeblink detection algorithm, six previously recorded EOG segments were analyzed. The segments are from two subjects "grc" and "the." The detection accuracy is determined by comparing the algorithm detection results with human expert's judgement. The performance of the WAM system is also included here for comparison. Tables 1 and 2 summarize the results based on experts I and II. We note that data sets 'the1" and 'the2" are very difficult to analyze since the subject was instructed to create artifacts such as excessive head movement, squinting, talking, swallowing, and horizontal eye movement during the recording.

To obtain the results in Table 1, both detection algorithms were trained on the initial data segment of the EOG file. The training of the WAM system requires the human expert to manipulate the system parameters manually while the training of the new eyeblink detection algorithm is done automatically. Also the amount of data used for training the WAM system is larger than the amount for the new EOG analysis algorithm. As one can see from the table that, except for one segment, the new eyeblink detection algorithm performed better or at least as good as the the WAM system, according to this particular expert.

For the results reported in Table 2, the training is done only once for each subject. The performance is significantly better for the new algorithm comparing to the WAM system. We note that based on the number of total blinks given by the two experts, expert II tends to have a more rigid standard for classifying an event as an eyeblink.

1.6 Conclusions

The new eyeblink detection algorithm is superior at least in the following four aspects when compared to the existing eyeblink detection algorithm on the WAM system:

1. *Automatic training.* No human operator is needed in the training phase. This feature allows real-time operation of the eyeblink detection process and maintains the consistence of the eyeblink detection criteria.

2. *Intra-subject invariability.* Data analysis results suggest that the new algorithm is more robust to the intra-subject eyeblink pattern variability. The new detection algorithm only needs to be trained once for each subject, then it can accurately detect eyeblinks from any EOG measurements made for that subject.

3. *Robust to artifacts.* The new algorithm incorporates nonlinear signal processing methods to reduce the effects of artifacts and recording noises. Analysis results confirm the effectiveness of data pre-processing. This is in contrast to the WAM system which often fails when there are artifacts in the EOG.

4. *Precise description.* The new algorithm assigns a likelihood value to each possible eyeblink event. This classification scheme is in parallel to the judgement of an expert thus makes it possible to label each eyeblink event (e.g., a blink, a small blink, a possible blink) more precisely.

PART II: EEG Study

2.1 Introduction

Electroencephalogram (EEG) is a direct result of brain's electric activity. Since its first recording by Berger [5], many research efforts have been devoted to establish a definite connection between the EEG and states of the brain. This report summarizes some preliminary results of the attempt to develop reliable indicators of mental workload conditions based on EEG measurements. A real-time cognitive state monitoring system based on EEG would have many important applications ranging from operator monitoring to interaction and control of man/machine systems [6].

EEGs have been successfully used to classify brain injuries [7], neurologic and psychiatric disorders [9], effects of drugs [10]. However, the reliability and accuracy of using EEGs to discriminate various mental activities/conditions have not reached a satisfactory level. For example, there is still no practically reliable means to differentiate between levels of mental workload based on EEGs. The shortcoming may be due to that the underlying EEG variations caused by mental workload changes are not as distinct as in other cases such as brain injuries. To be able to detect the subtle changes, EEG analysis algorithms must be of high sensitivity. Unfortunately, most analyses are non-parametric (e.g., short-term FFT spectrum-based) and they are susceptible to noise and other random fluctuations. Thus the variability in the analysis results cannot be reliably interpreted and related to the mental workload condition changes. A parametric model-based method is proposed and developed for accurate quantification of changes in EEG associated with mental workload variations.

2.2 Autoregressive Modeling of EEG

Although EEG is the output of a very complex nonlinear system, it has been successfully modeled as an autoregressive (AR) process ([11]–[14]):

$$x[k] = \sum_{i=1}^{P} a_i x[k-i] + e[k], \tag{5}$$

where P is the order of the AR process and $e[k]$ is the unpredictable part of $x[k]$ (residue error). It is reasonable to expect that when the condition of the system (the brain) generating the EEG changes (e.g., operates under different mental workload conditions), the AR model which best describes the EEG would also change. So if we can detect the changes in EEG via its AR model parameters, we can infer the changes occurring in the brain (e.g., a different level of mental workload). We hypothesize that the parametric model will be more sensitive to the underlying changes in EEG than traditional FFT spectra. Therefore, the AR model would be a better alternative for detecting workload condition variations.

To determine the AR model for a segment of EEG, the order P must be determined first. Then the parameter set $\{a_i, i = 1, \cdots, P\}$ needs to be estimated. Finally, a diagnostic checking should be performed to see if the residue error $e[k]$ is indeed white. Our initial effort concentrated on how to develop a distance measure based on AR model parameters to quantify the EEG changes related to the workload variations. Based on a priori knowledge, we choose $P = 6$ as the model order for all EEG segments in this study. It is possible that variations in optimal AR model order for EEG segments carry information which might be useful for workload condition classification and the issue will be dealt with in the future.

Given the optimal AR model order, we can now give an intuitive explanation to the question that why the parametric model is a better alternative to non-parametric methods like short-term FFT. Consider a segment of EEG of length 256, a short-term FFT analysis would give 128 individual frequency components. On the other hand, 6 AR model parameters are determined by the same 256 points. It is not difficult to see that the estimation variances of the 6 AR model parameters are smaller than the variances for the 128 FFT spectrum components when both sets of estimates are based on the same amount of data. In other words, any variations in the AR model parameters can be attributed to the underlying changes in EEG with greater confidence than the same variations in the FFT spectrum components.

In many instances, EEG spectrum is needed for visual inspection or other analysis. It is worth pointing out that the desired spectrum information can still be obtained from the AR model parameters via the following frequency response formula:

$$H(e^{j\Omega}) = \frac{1}{1 - \sum_{n=1}^{P} a_n e^{-jn\Omega}}, \quad 0 \leq \Omega \leq 2\pi. \tag{6}$$

However, the AR model parametric approach is not without any drawbacks. The primary liability is the AR model assumption itself. If the time series (EEG signal here) should not be modeled by an AR process, any conclusion we draw based on the AR assumption may be flawed. Since many research work on EEG analysis has demonstrated the appropriateness and effectiveness of the AR modeling assumption, it is worthwhile to explore the feasibility of the AR model approach for mental workload classification.

With the assumption that each EEG segment can be modeled as an AR process of order P, one needs to find the parameter set $\{a_i, i = 1, \cdots, P\}$ in (5) in order to completely specify the model. To determine the appropriateness of a set of AR parameters $\{a_i, i = 1, \cdots, P\}$, one often resorts to the following mean square error (MSE) criterion

$$\mathcal{E} = E[(x[k] - \sum_{i=1}^{P} a_i x[k-i])^2]. \tag{7}$$

The set $\{a_i, i = 1, \cdots, P\}$ which minimizes $\mathcal{E}$ is considered to be optimal.

From (7), the following normal equations can be obtained

$$\begin{pmatrix} r_x(0) & r_x(1) & \cdots & r_x(P-1) \\ r_x(1) & r_x(0) & \cdots & r_x(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(P-1) & r_x(P-2) & \cdots & r_x(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{pmatrix} = - \begin{pmatrix} r_x(1) \\ r_x(2) \\ \vdots \\ r_x(P) \end{pmatrix} \tag{8}$$

where $r_x(j)$ is the autocorrelation function of $x[k]$ defined as $r_x(j) = E(x[k]x[k-j])$. One simple and straightforward way of estimating the autocorrelation function $r_x(j)$ is through short-term averaging. Once the values of $r_x(j)$ become available, the normal equations can be solved for $[a_1 \, a_2 \, \cdots \, a_P]^T$. Various algorithms have been developed [15] to efficiently estimate the autocorrelation function $r_x(j)$ and to solve the normal equations (8). In this project, we will utilize Levinson's method for solving (8) [16].

2.3   Itakura Distance

Given two segments of EEG, one would like to have a simple, compact, and efficient descriptor to measure the degree of similarity or dissimilarity (distance) between them. Such a descriptor is very important for detecting subtle changes in EEG and for classifying mental workload conditions. Suppose two sets of AR parameters are obtained for two EEG segments (say one $x[k]$ is from baseline workload and the other $y[k]$ is from high workload):

$$x[k] = \sum_{i=1}^{P} a_i x[k-i] + e_x[k] \tag{9a}$$

$$y[k] = \sum_{i=1}^{P} b_i y[k-i] + e_y[k]. \tag{9b}$$

Since the AR model parameters describe the EEG segments, it is reasonable to expect that the distance between $x[k]$ and $y[k]$ can be reflected in a proper distance measure between the two sets of AR parameters.

The question of how to measure the distance between two AR processes has been studied extensively in digital speech processing in applications like objective speech quality measurement and speech recognition [17]. One of the most successful methods for measuring the distance between two speech utterances modeled as AR processes is the Itakura distance measure [18]. In EEG analysis, the Itakura distance has been applied successfully to detect EEG changes related to brain injuries [7].

The Itakura distance can be explained intuitively as follows: Let the optimal AR model parameter set for the signal $x[k]$ be denoted as

$$\alpha = [1 \; -a_1 \; -a_2 \; \cdots \; -a_P]^T.$$

Recall that the optimal parameter set α minimizes energy of the residue error $e_x^\alpha[k]$:

$$e_x^\alpha[k] = x[k] - \sum_{i=1}^{P} a_i x[k-i]$$

and let the minimum energy be denoted as ξ_α. It is easy to show that the minimum energy is given by $\alpha^T R_x \alpha$. Now let the signal $x[k]$ pass through a filter with the parameter set

$$\beta = [1 \; -b_1 \; -b_2 \; \cdots \; -b_P]^T$$

or equivalently

$$e_x^\beta[k] = x[k] - \sum_{i=1}^{P} b_i x[k-i]. \tag{10}$$

The energy of this error output $e_x^\beta[k]$ is denoted as ξ_β and is given by $\beta^T R_x \beta$. The closer to α the parameter set β is, the smaller the error energy ξ_β will be (but always $\geq \xi_\alpha$). This is true because by definition α minimizes the error energy ξ. Thus, a comparison between the error energy ξ_α and ξ_β becomes an effective indicator of how close the two sets of AR parameters are.

Formally, the Itakura distance between β and α is defined as

$$d_I'(\alpha, \beta) = \log \frac{\xi_\alpha}{\xi_\beta} = \log \frac{\beta^T R_x \beta}{\alpha^T R_x \alpha} \tag{11}$$

where matrix R_x is the correlation matrix for $x[n]$

$$R_x = \begin{pmatrix} r_x(0) & r_x(1) & \cdots & r_x(P) \\ r_x(1) & r_x(0) & \cdots & r_x(P-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(P) & r_x(P-1) & \cdots & r_x(0) \end{pmatrix}.$$

A distance close to zero indicates an energy ratio close to one, thus a close match between the two EEG segment models.

The parameter sets α can also be tested to see how well it models the EEG signal $y[k]$ (with the optimal AR model parameter set β) by calculating

$$d_I'(\beta, \alpha) = \log \frac{\alpha^T R_y \alpha}{\beta^T R_y \beta}. \tag{12}$$

A symmetric distance measure can be defined by combining (11) and (12)

$$d_I(\alpha, \beta) = \frac{1}{2}(d_I'(\alpha, \beta) + d_I'(\beta, \alpha)). \tag{13}$$

The Itakura distance defined above is used to analyze EEG for mental workload classification. The EEGs were recorded when subjects were performing tasks of various difficulty levels (e.g., communications only,

low workload, medium workload, and high workload). The EEGs were sampled at 200 Hz. The objective is to discriminate different mental workload conditions based on the EEG analysis.

The concept of coherence has been used in EEG analysis for various purposes [8]. Here we point out the conceptual difference between coherence and Itakura distance: Coherence measures the degree of linear association between two segments of EEG while Itakura distance reveals the difference between linear models of two EEG segments. For example, if two EEG segments are modeled by two distinct linear AR models, the Itakura distance would be greater than zero since the two AR models are different. On the other hand, the coherence between the two EEG segments can still be perfect (equal to one) if these two AR models are both invertible. The Itakura distance is therefore a much more sensitive indicator of changes in EEG (both linear and nonlinear) and the coherence is only capable of detecting nonlinear changes in EEG.

2.4 Data Analysis Results

Two types of EEG analysis results are presented here. Temporal distances are used to discriminate EEGs from the same recording site but under different mental workload conditions. Spatial distances attempt to quantify changes in the relations of EEGs between different recording sites as the mental workload condition varies.

2.4.1 Temporal Distance

The Itakura distance was first used to measure the distance between EEG segments from low, medium, high workload and EEG segments of (communications only) baseline. That is, if $x[k]$ is the EEGs from baseline workload condition and $y_{l,m,h}[k]$ are the EEGs from low, medium, and high workload conditions, we hope to reveal that the dissimilarity, as measured by the Itakura distance, between the EEG segments from high workload and baseline would be greater than that between the EEG segments from low workload and baseline conditions. It follows that the following relation should hold:

$$d_I(\alpha, \beta_{low}) < d_I(\alpha, \beta_{medium}) < d_I(\alpha, \beta_{high}). \tag{14}$$

Analysis results so far did not reveal any clearly identifiable patterns for different mental workload conditions based on the temporal distance measure. One possible explanation due to nonstationarity of the EEG is suggested as follows. The characteristics of EEGs normally varies with time even under the same general workload condition. Therefore the temporal distance between a segment from baseline and another segment from high workload is results of possible EEG variations caused by workload condition changes as well as the nonstationarity of the EEG within the each workload. It is thus necessary to discriminate the effects of EEG nonstationarity within each workload condition and EEG variations across workload conditions.

2.4.2 Spatial Distance

The spatial relations between EEGs from different regions of the brain were then studied, partly motivated by the difficulty encountered in temporal relation study. It was hypothesized that as the mental workload condition changes, the relations between EEGs from different regions of the brain would also change. We tested this hypothesis by analyzing 57 channels of EEG available which were recorded from a total of ten subjects under four different mental workload conditions (baseline, low, medium, and high). The Itakura distance was used to quantify the pair-wise comparison of similarity (or dissimilarity) of all 57 channel EEGs under all four mental workload conditions. Since the distance measure was done with respect to EEGs recorded from different regions but at the same time, the effects of the EEG nonstationarity should be minimum.

The EEGs were divided into non-overlapping 256 point segments and the Itakura distance between all recording electrode pairs ($57 \times 56/2$) were calculated for each 1.28 second (256 point) segment. The Itakura distances were then averaged over ten segments (for a total of 12.8 seconds). Figures 4 and 5 show two samples of such distance map. The square marked with an "R" indicates the reference electrode position
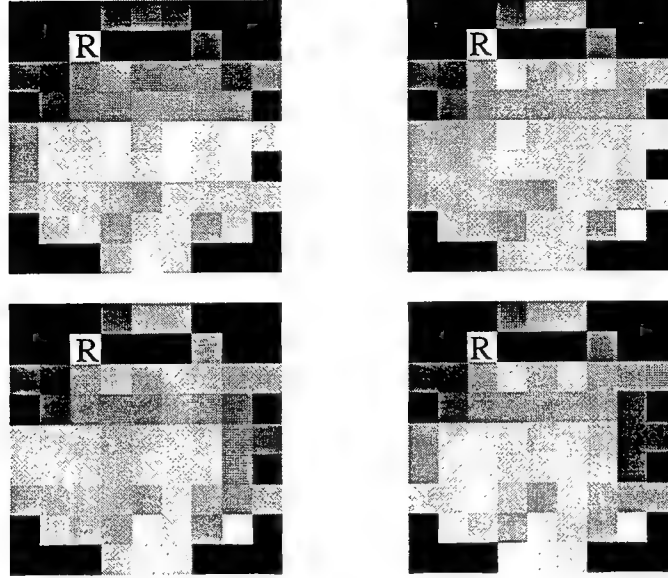
Figure 4: Itakura distance maps for four different mental workload conditions. Clockwise from top-left corner: baseline communications only, low workload, high workload, and medium workload. A darker color represents a smaller distance. Reference electrode (marked with "R") is F3A.

and the darkness of each square represents the Itakura distance between the EEG segment from that electrode and the EEG segment from the reference electrode. A darker square indicates a smaller Itakura distance. We note that information presented in each distance map is enormous and we have 57 such maps, one for each reference recording site.

In order to be able to visually extract meaningful information from the distance maps, we limited the display to the five electrode positions whose EEGs have the minimum distance to the reference electrode EEG. Figures 6 and 7 are two such examples. Normally, one expects EEGs from neighboring electrodes to be similar to each other (thus a small Itakura distance) and as the physical distance of the electrodes increases, the Itakura distance of the corresponding EEGs will also increase. Consequently, the five minimum distance electrode positions should be physically close to the reference electrode position. Indeed, this is true for the baseline communications-only workload condition. At the high workload condition for reference electrode F3A (Figure 6, lower right corner), we observe a distinct EEG distance pattern: EEGs from electrodes physically apart from each other exhibit a high degree of similarity as measured by the Itakura distance. Furthermore, the frequency at which the distinct EEG distance pattern occurs increases as the workload level increases. We note that the distinct distance pattern only happens for a few specific reference electrodes.

In order to numerically summarize the distinct distance pattern, we developed the following workload index for a given reference electrode position r

$$W(r) = \frac{\displaystyle\sum_{e \in E(r)} (D - d_I(e,r))d_P(e,r)}{\displaystyle\sum_{e \in \{\text{all electrodes}\}} d_P(e,r)} \tag{15}$$

where D is a constant no less than the largest d_I, $E(r)$ is the set of electrodes e whose EEG Itakura distance $d_I(e,r)$ to reference electrode r is among the five smallest, and $d_P(e,r)$ is the physical distance between electrodes e and r. The denominator normalizes the index. A higher index value would correspond to a distance map where EEGs from far-apart electrode pairs exhibit a close association as measured by the

Figure 5: Itakura distance maps for four different mental workload conditions. Clockwise from top-left corner: baseline communications only, low workload, high workload, and medium workload. A darker color represents a smaller distance. Reference electrode (marked with "R") is T3L.

Figure 6: Itakura distance maps for four different mental workload conditions. Clockwise from top-left corner: baseline communications only, low workload, high workload, and medium workload. The five light gray squares represent the electrode sites whose EEGs are closest to the reference EEG as measured by the Itakura distance. Reference electrode (marked as "R") is F3A.
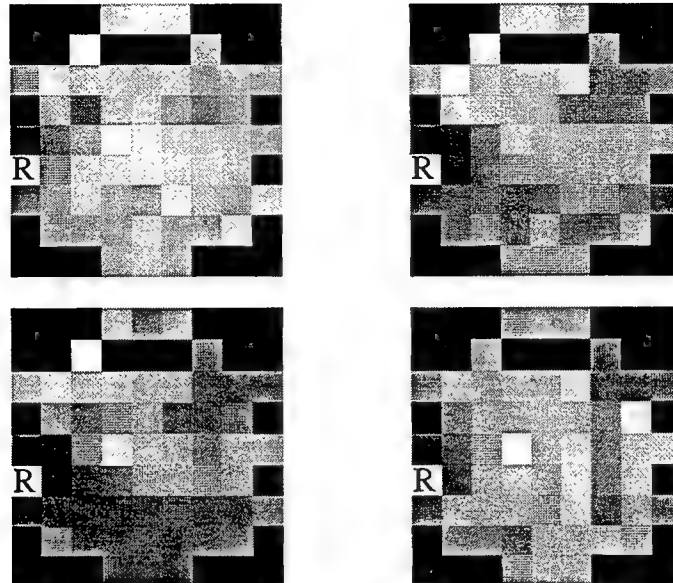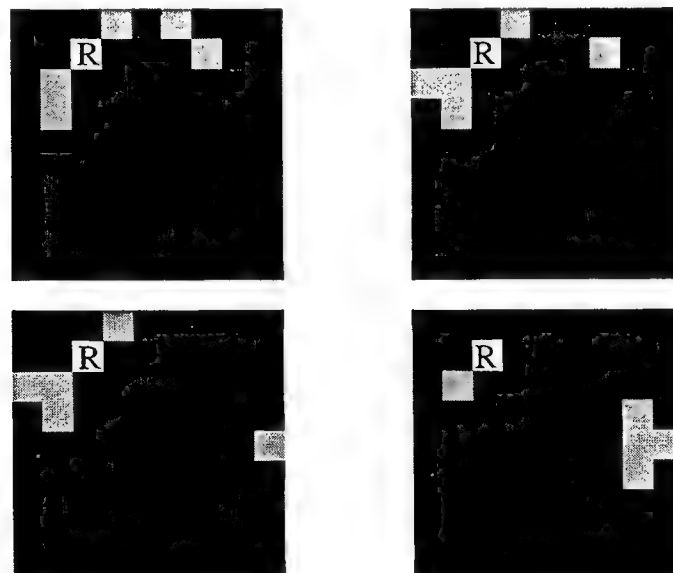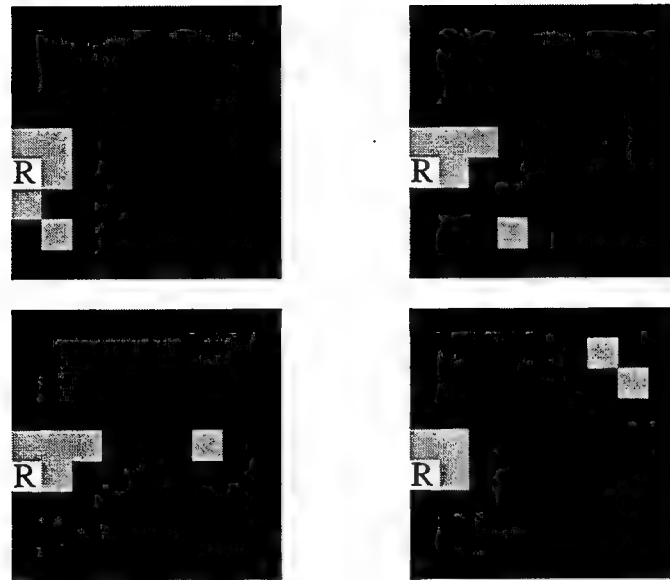
Figure 7: Itakura distance maps for four different mental workload conditions. Clockwise from top-left corner: baseline communications only, low workload, high workload, and medium workload. The five light gray squares represent the electrode sites whose EEGs are closest to the reference EEG as measured by the Itakura distance. Reference electrode (marked as "R") is T3L.

Itakura distance. Based on the discussion we have, EEGs from higher workload conditions are more likely to have this property. Therefore, this simple-minded index is expected to give us an indication of workload conditions.

Figure 8 shows sample distance maps for all four workload conditions (from top to bottom: baseline communications only, low workload, medium workload, and high workload) for a 180-second period (from left to right, each frame corresponds to 12.8 second EEG). The number at the top of each distance map is the workload index value for that map. As suggested before, the index values are generally larger for higher workload conditions than the values for lower workload conditions.

Table 3 summarizes the workload index values of several selected reference electrodes for nine separate trials, three for each workload condition. The index values are the averaged ones over a 180-second duration with each segment duration being 12.8 seconds. As one can see from the table, the relative sizes of workload index are in good agreement with the mental workload conditions. The absolute boundary varies with electrode sites for discriminating workload conditions. Though the workload index values fall into two distinct sets for low and high workload conditions, the workload index for medium workload may have values in either low or high workload regions (e.g., electrode O2). This ambiguity is not unique to workload index analysis, it was also observed in other studies like performance analyses and subject's self-evaluations [26].

Finally, three low workload trials are randomly paired with three high workload trials and the workload conditions are classified for each 12.8 second segment. The classification results are listed in Table 4. The classification rate of only 70–80 % suggest that when the brain is challenged with tasks of a greater difficulty, it may not necessarily be in a higher mental workload condition at all time even though the brain may be in that higher mental workload condition on average, as measured by the grand average of workload index in Table 3.

Figure 8: Sequence of Itakura distance maps under four different mental workload conditions. Reference eletrode site is indicated by a white square (F3A). The five electrode sites whose EEGs are closest to the reference EEG are marked as light gray. From top to bottom: baseline communications only, low workload, medium workload, and high workload. From left to right is a 180-second time sequence with each frame corresponding to 12.8 second EEG. The number at the top of each distance map is the workload index value for that map. The workload index values are larger for higher workload conditions than the values for lower workload conditions.

electrode	low			medium			high		
site	trial 6	trial 11	trial 18	trial 8	trial 12	trial 19	trial 5	trial 14	trial 17
F3A	28.71	25.99	26.02	39.46	33.59	32.55	38.99	61.95	45.65
F6	17.91	16.72	15.26	17.24	20.80	19.89	41.14	18.98	24.45
FP1	23.56	19.57	21.84	27.66	23.81	22.91	32.87	39.67	25.32
C2A	20.30	21.98	20.86	23.21	20.78	24.47	37.07	29.81	25.82
T3L	18.74	18.06	15.02	22.70	27.35	31.17	39.95	31.17	18.74
O2	24.38	21.48	19.80	27.56	27.90	23.03	26.64	30.16	26.31

Table 3: Averaged workload index values of EEG from six selected electrode sites under three different mental workload conditions. The workload index values correspond to three mental workload conditions very well. Some low/medium and medium/high ambiguity observed here is consistent with other analyses such as performance analysis.

low–high pair	F3A	F6	FP1	C2A	T3L	O2	F4A	PZA	F5
trial 6–trail 5	86%	93%	79%	86%	100%	57%	93%	57%	71%
trial 11–trial 14	93%	43%	100%	86%	71%	79%	71%	79%	100%
trial 17–trial 18	79%	86%	71%	71%	57%	64%	86%	86%	71%
average	86%	74%	83%	81%	76%	67%	83%	83%	80%

Table 4: Classification results for three randomly paired low–high workload experiments based on workload index analysis of EEG. Classification is done once every 12.8 seconds and the entries in the table are the correct classification rate.

2.5 Future Research Directions

The results obtained based on spatial distance are encouraging and they warrant further investigations. Many aspects of the spatial distance approach need to be studied before the appropriateness of this method can be determined for mental workload classification. Future research undertakings are suggested as follows to improve the reliability and accuracy of the mental workload classification based on the EEG analysis:

1. *EEG segmentation.* A better segmentation scheme may lead to a more accurate quantification of the EEG changes due to mental workload variations. Current study uses a fixed non-overlapping segment length of 256 points. Other segmentation length (with and without overlaps) should be investigated in order to determine the optimum segmentation length.

   Under the current protocol, all workload conditions involve the same set of tasks and a particular workload condition is determined by the frequency at which the subject is required to perform the tasks. Instead of segmenting the EEG on a fixed time interval, it may be advantageous to segment the EEG based on the tasks being performed. Such a segmentation scheme allows us to investigate the EEG pattern changes under various workload conditions without the influences of task difference in a given EEG segment.

2. *AR model order.* Optimal AR model order carries important information about the characteristics of a time series. Optimal AR model order of EEG was shown to be a good indicator of possible brain injury [25]. It is reasonable to expect that the optimal AR model order may also contribute significantly to the classification of the mental workload conditions.

3. *Band-based AR model.* Traditionally, EEGs were analyzed in terms of well-defined frequency bands. It may be helpful to filter the original EEG signal to obtain bandpassed EEG signals according to the frequency ranges of δ, θ, α, and β bands. An AR model is then established for each bandpassed EEG. The Itakura distance in this case allows us to compare the EEG segments within a particular frequency range, instead of comparing them over the whole spectrum.

4. *Neural network as classifier.* Our objective is to reliably classify mental workload conditions based on EEG analyses. The AR model provides us with an accurate and efficient description of the EEG signal and the Itakura distance helps us summarize the temporal or spatial relations between the EEGs. The distance maps in Figures 4 –7 display such information visually. And the workload index proposed in (15) represents a first attempt to utilize the enormous amount of information in the distance map for workload classification.

   However, it is very likely that the relation defined in (15) cannot accurately map a 56-dimensional vector space (56 distance value) to a one dimensional vector space (workload conditions). Unless the sample points in the 56-dimension space for different classes (different workload conditions) are well separated, the desired mapping must be highly nonlinear. Thus, to improve the accuracy of the classification based on distance patterns, a mapping more sophisticated than the workload index in (15) should be developed. A neural network of appropriate size may serve the purpose very well as

it possesses many desired properties: 1) it can approximate any nonlinear mappings with arbitrary accuracy; 2) it can tolerate some inconsistency in the training data; and 3) it does not require a priori knowledge of the nonlinear mapping to be approximated. Neural network classifiers will be investigated to evaluate their performance of classifying workload classes based on the Itakura distances and other EEG features which may be developed in the future.

5. *Other relevant EEG features.* Although the Itakura distance was shown to be effective in detecting EEG changes related to workload conditions, the effectiveness of other distance measures should also be investigated and compared for mental workload condition classifications based on EEG analysis. Other popular AR model-based methods include spectrum distance and cepstrum distance [17]. The same neural network structures developed for the Itakura distance can be used to classify workload conditions based on these new features and the classification results can be compared to those of the Itakura distance.

Other features which are not based on the AR model assumption deserve our attentions as well. For example, Kullback-Leibler information [23] was successfully used to discriminate EEGs from a normal brain and an injured one [7]. Again the information measure obtained can be fed to the neural network for workload classification.

References

[1] J. Veltman and A. Gailard, *Biological Psychology*, vol. 42, pp. 323–342, 1996.

[2] G.F. Wilson and F. Fisher, "The use of cardiac and eye blink measures to determine flight segment in F4 crews," *Aviat. Space Environ. Med.*, vol. 62, pp. 959–962, 1991.

[3] J. Stern and D. Dunham, "The Ocular System", in *Principles of Psychophysiology*, edited by John Cacioppo and Louis Tassingary , Cambridge University Press, New York, NY, 1990.

[4] R.W. Hamming, *Digital Filters*, 3rd edition, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[5] H. Berger, "Öber das elektroenzephalogramm des menschen," *Arch Psychiat Nervenkrankheiten*, vol. 87, pp. 527–570, 1929.

[6] Glenn F. Wilson and Frank Fisher, "Cognitive task classification based upon topographic EEG data," *Biological Psychology*, vol. 40, pp. 239–250, 1995.

[7] X. Kong, V. Goel, and N. V. Thakor, "Quantification of injury-related EEG signal changes using Itakura distance measure", in *Proc. of Int. Conf. on Acoustics, Speech, Signal Processing*, Detroit, MI, May 1995, pp. 2947–2950.

[8] N.V. Thakor, X. Kong, and D.F. Hanley, "Nonlinear changes in brain's response in the event of injury as detected by adaptive coherence estimation of evoked potentials," *IEEE Trans. Biomedical Engineering*, vol. 42, pp. 42–51, 1995.

[9] E.R. John, L.S. Prichep, J. Fridman, and P. Easton, "Neurometrics: Computer-assisted differential diagnosis of brain dysfunctions," *Science*, vol. 239, pp. 162–169, 1988.

[10] W.M. Herrmann, "Development and critical evaluation of an objective procedure for the electroencephalographic classification of psychotropic drugs," in *EEG in Drug Research*, pp. 249–351, Gustav Fisher, New York, NY, 1982.

[11] G. Bodenstein and H. Praetorius, "Feature extraction from electroencephalogram by adaptive segmentation", *Proc. of IEEE*, vol. 65, pp. 642–652, 1977.

[12] B. Jansen, J. Bourne, and J. Ward, "Autoregressive estimation of short segment spectra for computerized EEG analysis", *IEEE Trans. Biomed. Eng.*, vol. 28, pp. 630–638, 1981.

[13] L. Zetterberg, "Estimation of parameters for a linear difference equation with application to EEG analysis", *Math. Biosci.*, vol. 5, pp. 227–275, 1969.

[14] H. Steinberg, T. Gasser, and J. Franke, "Fitting autoregressive models to EEG time series: An empirical comparison of estimates of the order", *IEEE Trans. on Acoustics, Speech, Signal Processing*, vol. 33, no. 1, pp. 143–150, 1985.

[15] J. Makhoul, "Linear prediction: A tutorial review", *Proc. of IEEE*, vol. 63, pp. 99–118, 1975.

[16] *MATLAB User's Guide*, MathWorks Inc., 1992.

[17] J. Deller Jr., J. Proakis, and J. Hansen, *Discrete-Time Processing of Speech Signals*, Macmillan, New York, NY, 1993.

[18] F. Itakura, "Minimum prediction residual principle applied to speech recognition", *IEEE Trans. on Acoustics, Speech, Signal Processing*, vol. 23, pp. 67–72, 1975.

[19] H. Akaike, "Fitting autoregressive models for prediction", *Ann. Inst. Statist. Math.*, vol. 20, pp. 243–247, 1969.

[20] J. Rissanen, "Modeling by shortest data description", *Automatica*, vol. 14, pp. 465–471, 1978.

[21] W. Gersch, "Spectral analysis of EEG's by autoregressive decomposition of time series", *Math. Biosci.*, vol. 7, pp. 205–222, 1970.

[22] A. Isaksson, A. Wennberg, and L. Zetterberg, "Computer analysis of EEG signals with parametric models", *Proc. of IEEE*, vol. 69, pp. 451–461, 1981.

[23] S. Kullback and R. Leibler, "On information and sufficiency", *Annals of Mathematical Statistics*, vol. 22, pp. 79–86, 1951.

[24] J. McEwen and G. Anderson, "Modeling the stationarity and Gaussianity of spontaneous electroencephalographic activity", *IEEE Trans. on Biomed. Eng.*, vol. 22, pp. 361–369, 1975.

[25] X. Kong, V. Goel, and N.V. Thakor, "Quantification of injury-related EEG signal changes using distance and information measures," submitted to *IEEE Trans. Biomedical Engineering*.

[26] Glenn F. Wilson, *personal communications*, May–August, 1996.

PRELIMINARY STUDIES OF HUMAN ELECTROENCEPHALOGRAM (EEG) CORRELATES OF Gz ACCELERATION TOLERANCE

Charles S. Lessard
Associate Professor

Bioengineering Program,
Industrial Engineering Department
Texas A&M University
College Station, TX 77843

PRELIMINARY STUDIES OF HUMAN ELECTROENCEPHALOGRAM (EEG) CORRELATES OF Gz ACCELERATION TOLERANCE

Charles S. Lessard
Associate Professor
Bioengineering Program
Texas A&M University

Abstract

This report presents preliminary findings of two separate studies whose goal is the development of a system for detection of acceleration (+Gz) induced loss of consciousness (G-LOC) in pilots of high performance fighters. The purpose of the first study was to evaluate a helmet mounted system; presently under development by the Israeli Air Force for determination or detection of Gz induced loss of consciousness (G-LOC). The results show large amounts of delta activity in the EEG; however, the dry EEG system has a greater percentage of delta when compared to the wet EEG system. Of more interest is the increase of high frequencies in the Beta bands (Beta 3 to Beta 6 or from 24 to 36 Hz), as the G-level increases to 9-Gs. The high frequency increase is greater and easier to note in the wet EEG data than in the dry EEG data.

For the second study, the subjects were in different stages of training; thus, the acceleration profile for any subject is based on the individual's training requirement. The acceleration-profiles in this study could include any set of rapid onsets to 2-Gz, 4.5-Gz, 5-Gz, 7-Gz, 8-Gz and/or 9-Gz for varying duration of time. More advanced trainees experience the Standard Air Combat Maneuver (SACM). In this study, it was observed that the experienced rider's (subject A) anti-G countering maneuver during the SACM showed maximal EEG activity at the first +7-G encounter. The maximal anti-G countering strain resulted in high levels of physiological electrical activity in the 16 to 32 Hz range. Duing subsequent +7-Gz accelerations, less EEG activity was observed in the 12 to 28 HZ range. The results from the inexperienced centrifuge rider (subject B) indicate unusually high activity during the 15 seconds of sustained (constant) Gz-levels of 5 and show elevated EEG activities in the 12 to 20 Hz range during 5-Gz acceleration and Beta 1 (12-16 Hz) during the 7-Gz acceleration. An evaluation of the EEG spectra at 7-Gz shows a very high peak at 13.6 Hz on all EEG channels; however, it is most prominent in the parietal area (40% of spectral power) as compared to the occipital area (34% of spectral power).

23-2

PRELIMINARY STUDIES OF HUMAN ELECTROENCEPHALOGRAM (EEG) CORRELATES OF Gz ACCELERATION TOLERANCE

Charles S. Lessard

Introduction

This report presents preliminary findings of two separate studies whose goal is the development of a system for detection of acceleration (+Gz) induced loss of consciousness (G-LOC) in pilots of high performance fighters. There is abundant literature on the effects of high +G-levels on the performance of pilots, but there is very little written on the subject of central nervous system response, i.e., changes in the brain's electrical activities, at higher +Gz-levels. In addition, there are reports from animal studies (P. Werchan used rats, Burns and Werchan {1} used baboons) which indicated an increase of delta activity (2 seconds) preceding G-LOC of the animal, but no evidence exists that a similar precursor to G-LOC occurs in humans trained to counter the loss of consciousness. Part of the problem in the lack of data from humans lies in the acquisition of what is considered reliable, noise-free electroencephalograms (EEG) from the centrifuge environment.

Background on the Effects of +Gz Stress

Gillingham indicated that primary effects of increased acceleration (+Gz) are the loss of blood pressure from the cerebral circulation and decreased cardiac output. {2} These effects attribute to the possible loss of consciousness as the pilot engages in high +G levels during aerial combat maneuvers. Various devices and schemes have been developed to counter the adverse effects of high +Gz levels. For example, the anti-G suit with rapidly opening valves and positive pressure breathing have proved effective in improving the +Gz tolerance by 3 to 5 Gs. In addition, pilots are being trained to counter the +Gz stress with the "L-1" anti-G straining maneuver (AGSM). The L-1 maneuver requires the pilot to vigorously contract skeletal muscles (extremities first), inhale, and increase the intrathoracic pressure by vigorous contraction of the abdominal muscles with forceful exhalation at the end of three seconds in order to reoxygenate the body. The L-1 maneuver is repeated to elevate and maintain high arterial blood pressure to the head and at the same time prevent blood from pooling in the lower extremities. The effectiveness of the L-1 AGSM assumes a normal or better physical condition with optimal physiological and biological responsiveness to the effects of +Gz-induced stresses.

FIRST STUDY

The purpose of this study is to evaluate a helmet mounted system; presently under development by the Israeli Air Force for determination or detection of Gz induced loss of consciousness (G-LOC). The G-LOC detection system, self-contained and built into a standard flight helmet, consists of a set of dry electrodes, and built-in preamplifiers, signal conditioning circuits, and analog-to-digital converter for acquisition of the electroencephalogram (EEG) and electro-oculogram (EOG) signals. Reports from animal studies indicate shifts or increases in the delta activity of the EEG immediately preceding G-LOC [1]; however, there is no scientific evidence to indicate that a similar increase in delta activity occurs in pilots who are actively countering the loss of consciousness. In fact, little information is available describing brain electrical activity changes at elevated Gz. This report compares the EEG activity acquired with the dry electrode system to EEG activity acquired with throw-away, electrocardiogram (ECG), paste electrodes (wet electrodes). Additionally, the EEG activity at different, constant Gz levels are compared and examined for possible trends in the EEG.

Subjects

Six volunteers were selected from fully qualified centrifuge subjects at the United States Air Force (USAF) Armstrong Laboratory. Each subject was individually fitted with a dry electrode helmet; placements of the dry electrodes were checked; and the sites for the wet electrodes were prepared, before attachment of the electrodes. After preparation, the wet electrode-skin impedance was measured. The helmets were prepared for each wearer based on 3-D skull topography to ensure a tight form fit.

Each subject experienced the same experimental protocol (acceleration profile) in a single session of one-hour duration. The first acceleration-profile is a rapid onset to 2-Gz followed by two-minutes at a constant 2-Gz level; subsequently, the subject is given three-minutes of rest at 1-Gz prior to any additional maneuver. The entire sequence of the acceleration profiles is as follows:

1. Rapid onset to 2-Gz level for 2-minutes: rest for 3-minutes at 1-Gz.

2. Gradual linear increase in acceleration from 1-Gz to 9-Gz level in a period of approximately 90-seconds.

3. Rapid deceleration from 9-Gz to 1-Gz level in 10-seconds: rest for 3-minutes at 1-Gz.

4. Rapid onset to 5-Gz level for 15 seconds: rest for 3-minutes at 1-Gz.

5. Rapid onset to 7-Gz level for 15 seconds: rest for 3-minutes at 1-Gz.

6. Rapid onset to 9-Gz level for 15 seconds: rest for 3-minutes at 1-Gz.

7. Standard air combat maneuver(SACM).

The SACM consists of repeated exposures to 4.5-Gz for 15-seconds followed immediately, by a rapid onset to 7-Gz for 15-seconds. The SACM sequence of changing Gz-level exposure is continued until either, the subject, medical monitor, or the centrifuge controller terminates the run. Termination of the sequence is based on any one of the following factors: subject reporting 100% loss of peripheral lights, subject reporting fatigue, subject loss of consciousness (G-LOC), or abnormal medical physiologic response.determined by the medical officer. During the experimental exposures to acceleration, each subject, wearing a standard G-suit (CSU-13-B/P), was in an upright seat with 13-degree tilt.

Five subjects wore oxygen masks which were not connected to any oxygen delivery system. One subject did not wear a mask. Placement of the reference electrode for the wet electrodes varied among the five subjects who wore oxygen masks, 3 on the cheek (Infra-orbital region) and 2 on the jaw hinge (styloid process). Of the three subjects whose reference electrode was secured on the right cheek bone below the center of the eye (styloid process), one subject's reference electrode was totally hidden by the oxygen mask. On another subject, the reference electrode was about 60% obscured by the oxygen mask. On the third subject, the reference electrode was totally visible and touching the oxygen mask. The reference electrodes placed on the right jaw-hinge (Infra-orbital region) of two subjects were completely obscured by the oxygen masks.

Data from one of the subjects could not be downloaded from the flash memory. Several runs from another subject could not be analyzed, because the output from the dry electrode system indicated saturation of the EEG amplifier.

Data Acquisition Method

The data acquisition method utilized for this study is as follows. Four dry electrodes (compressed silver-silver-chloride: Ag-Ag-Cl) and four wet electrodes are used to record the electroencephalograms (EEG) and electro-oculograms (EOG) from six volunteer subjects. Placements of the electrodes do not conform to the

standard 10-20 clinical EEG montage, but are based on the pilot's helmet fitting and available bare-skin contact surface for the dry electrode. Two dry electrodes are placed on the forehead without preparation or gel approximately one centimeter above the outer canthus of each eye, and two other dry electrodes are located about one centimeter below the outer canthus of the eyes on the flat structure of the zygomatic region.

The dry electrodes are mounted in the visor area of the helmet and are held in place by spring tension. The spring apparatus is adjustable and necessary to preclude contact between the dry electrodes and any part of the helmet or mask in order to avoid inducing motion artifacts in the EEG signals. The two dry electrodes on the forehead are used as the difference pair of active electrodes for acquiring a channel of differentially recorded EEG referenced to one of the dry electrodes on the cheek bone (zygomatic region). The non-reference dry electrode, on the other cheek bone, is used with the most distal electrode on the forehead to differentially record the EOG signal. Reference for the EOG signal is the same reference electrode used for EEG recording.

Three ECG recessed paste electrodes are used to record the EEG for comparison with recordings from the dry electrodes. Two wet electrodes are placed on the forehead above the eyebrow and as close to the dry electrodes as the circular adhering disk will permit. Reference for the wet electrodes is placed under the mask about one centimeter below the center of the subject's right eye. Two subjects had the reference wet electrode placed under the mask near the right jaw hinge (styloid process); unfortunately, this reference electrode site is over the masseter muscle that tends to increase electromyogram (EMG) activity in the EEG recording when the subject closes the glottis and performs straining maneuvers. Additionally physological movements of inhalation, exhalation, and talking appear to induce low frequency noise or motion artifacts in the EEG recordings.

The EEG signals from the dry electrodes are high-pass filtered at 0.5 hertz (Hz) and low-pass filtered at 30 Hz with a five-pole Butterworth filter. The dry electrode amplifier gain is set at a multipling factor of 10^4 or 10,000. While filter settings for EEG signals from the wet electrodes are the same as the dry electrodes, the gain appears (from post analysis) to be 10 times less. EEG and EOG signals from the dry and wet electrode systems are digitized at 128 samples per second with a 16-bit analog-to-digital (A/D) converter. The digital data are stored on to the flash memory of a laptop computer inside the centrifuge. The flash memory is downloaded between runs to other temporary data storage devices for subsequent analysis.

23-6

Three accelerometers are mounted on the pilot's helmet to record three orthogonal axis of acceleration (x, y, z,). The accelerometers are digitized at 32 samples per second. A total of eight channels are recorded with, 2-EEGs, 2-EOGs, 1-EMG, and 3-acceleration axis. Emphasis is placed on the comparison of dry and wet electrode EEG recordings during increased Gz-loading. Data from neither the EOG nor the rectified, integrated EMG are analyzed in this study.

Data Analysis

Digitized EEG data from the dry electrode system, the wet electrode system, and accelerometers (x, y, z axis) are imported into the digital signal processing software called "DADiSP" which stands for Data analysis and Display Software by the DSP Development Corporation, Cambridge, MA. The DADiSP program permits extraction of data segments from a data file; therefore, 15-second segments at constant acceleration are extracted from each subject's EEG file. The accelerometer data is used to determine which start and stop timing points are to be used in the extraction of a segment. Use of the accelerometer data assures that the EEG segment to be extracted is taken from the timing at a constant G-level or during the gradual acceleration to 9-G.

Subsequent processing of the data segments includes removal of the DC-component (DC-offset), calibration, scaling, and power spectral analysis via a mixed-radix Fast Fourier Transform (FFT) algorithm. Each EEG spectrum is exported as an ASCII, series data file for subsequent feature or parameter extraction. The resulting spectrum extends to 64 Hz, but only the frequencies from 1 to 40 Hz are of interest in the EEG analysis. The spectrum is divided into ten bands of four-Hz band width which correspond to the standard clinical EEG frequency bands, i.e., Delta (1-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta-1 (12-16 Hz), etc. to Beta 7 (36-40 Hz); thus, thirty parameters obtained from the spectrum include:

a. Total power in a band,

b. Average power in a band, and

c. Relative power in a band

Total power in a band is the summation of power within a clinical band. Average power in a band is the total power in a band divided by the number of spectral points in the band. Relative power in the band is the total power in the band divided by the total power in the entire spectrum. Relative power is the normalized spectrum and is

expressed in percentage (%)in the comparison charts.

In addition to analysis of the EEG data at constant Gz-levels, the gradual linear acceleration from 1-Gz to 9-Gz is evaluated using a three-dimension (3-D), spectral waterfall analysis and display. As in the previous analyses, the accelerometer channel is used to determine the extraction points. For this analysis, the y-axis accelerometer signal is used for the base in an overplot. The 90-seconds of the y-accelerometer segment and the same period of EEG segment are raveled or further subdivided into 30-subsegments. The power spectrum is calculated via the FFT for the first subsegment and the information is stored for subsequent plotting. A second spectrum is calculated with the last 75% of the data points in the first segment and the first 25% of the data points in the second segment. This is the equivalent of taking a window one-thirtieth (1/30) of the total data set and calculating a spectrum the size of the window as the window is moved one-fourth of its width over the entire data set. A total of 120 spectra are calculated and displayed on one 3-D plot with frequency on the x-axis, time on the y-axis, and magnitude of the spectra on the z-axis. In this study, the accelerometer segment is raveled first (30-segments with 75% overlay), the color is changed to red for ease of separation from the EEG , then the EEG is raveled (waterfall) and is overplotted in yellow on the accelerometer 3-D plot.

Results And Discussion Of First Study

Constant Level of Acceleration: The Microsoft EXCEL software package was used to reduce the high-resolution spectra into 10 EEG bands (4 Hz bandwidth) and 30 EEG parameters. EXCEL spreadsheets are used to organize the data for plots (3-D bar-charts) to show variation of clinical EEG bands with Gz-levels. The most useful set of parameters is the Relative Power, since these measures correspond to the normalized power spectra. In Figure 1, the average per cell is obtained from the four subjects whose data could be analyzed. The charts show large amounts of delta activity in the EEG; however, the dry EEG system (Figure 1 a.) has a greater percentage of delta when compared to the wet EEG system (Figure 1 b.). Of more interest is the increase of high frequencies in the Beta bands (Beta 3 to Beta 6 or from 24 to 36 Hz), as the G-level increases to 9-Gs. The high frequency increase is greater and easier to note in the wet EEG chart than in the dry EEG chart.

Additionally, a series of simple paired t-tests were run for a comparison of the signal comparability between the dry and wet EEGs. No difference between the electrical EEG signals from the dry and wet systems

Figure 1 a. 3-D Bar Charts show changes in relative power from dry electrode system of the clinical EEG frequency bands with increasing +Gz.
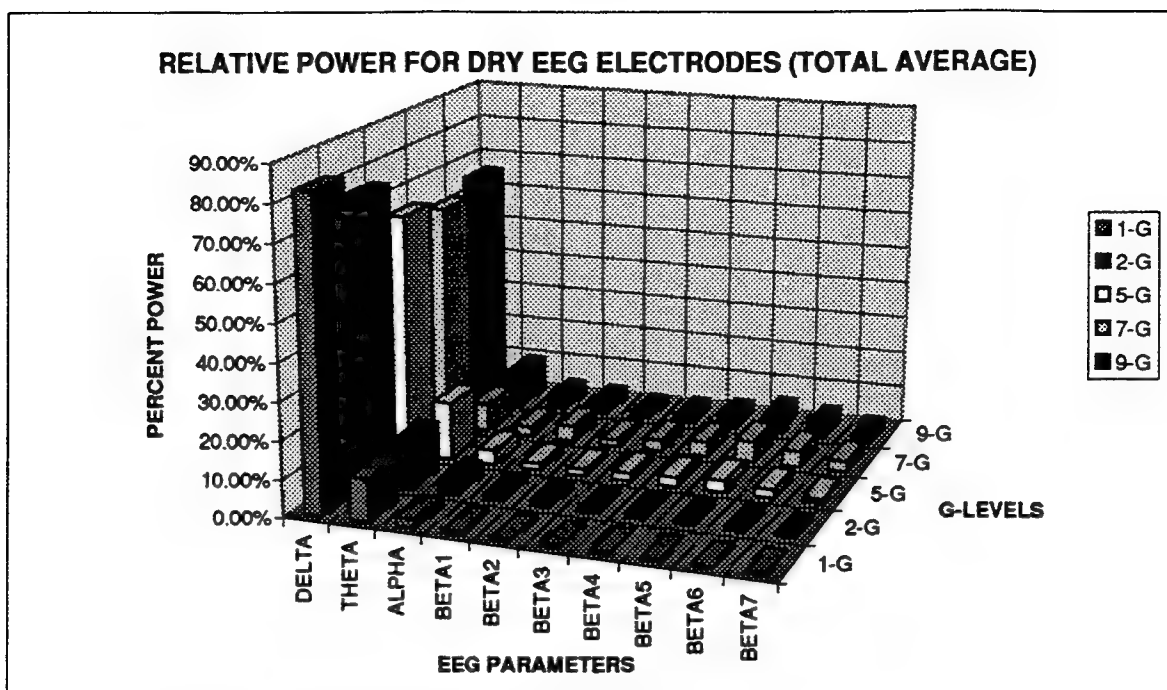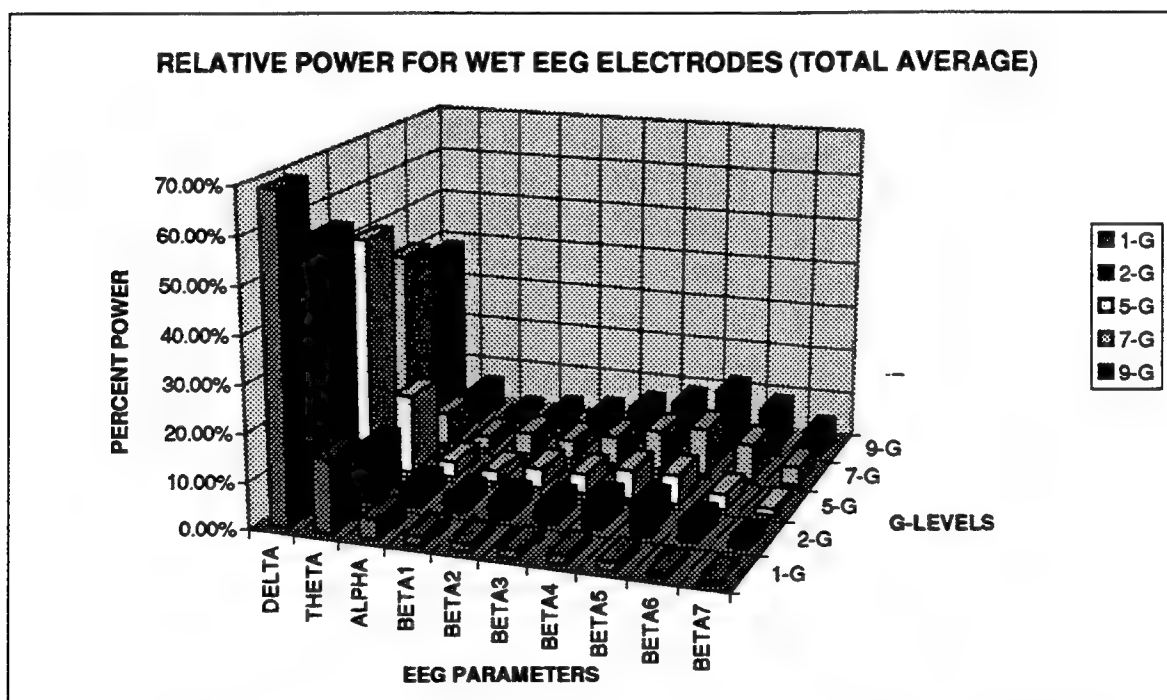

Figure 1 b. 3-D Bar Charts show changes in relative power of the clinical EEG frequency bands with increasing +Gz acceleration levels. Most noticable in 1 b is the increase in relative power of frequencies between Beta 3 and Beta 6 (24-36 Hz) as acceleration increases to +9 Gz. In figure 1 a, the dry electrode system shows a greater amount of Delta with less high frequency activity.

were predicted for the statistical comparison. The significance level for the paired t-test was set at alpha equal to 5%. In view of the small number of subjects, TABLE I is presented with the caution because, in most cases, the null hypothesis is "not rejected". TABLE I shows that the null hypothesis was rejected in favor of the alternate hypothesis that "there is significant difference in the EEG signals between the dry and wet systems," particularly for Beta-bands 4 through 7 at the 9-Gz level. This implies that the differences are large enough to be observed with a small sample.

TABLE I

RESULTS OF PAIRED t-TEST

Probability of Critical Value of T being less than the calculated t-statistics

Gz-LEVEL	1-Gz		2-Gz		5-Gz		7-Gz		9-Gz	
$\alpha = 0.05$ sig lev	1-tail	2-tail	1-tail	2-tail	1-tail	2-tail	1-tail	2-tail	1-tail	2-tail
EEG BANDS										
DELTA	-	-	.0395	.079	-	-	-	-	-	-
THETA	-	-	-	-	-	-	-	-	.039	.078
ALPHA	.0455	.091	-	-	-	-	-	-	-	-
BETA 1	-	-	.0195	.039	-	-	-	-	-	-
BETA 2	-	-	.034	.067	-	-	-	-	-	-
BETA 3	-	-	.048	.096	-	-	-	-	-	-
BETA 4	-	-	-	-	-	-	-	-	.0375	.075
BETA 5	-	-	-	-	-	-	-	-	.0105	.021
BETA 6	-	-	-	-	-	-	-	-	.024	.048
BETA 7	-	-	.049	.098	-	-	-	-	.011	.022

NOTE: Level of significance; $\alpha = 0.05$.

Gradual 1-Gz to 9-Gz Acceleration: The DADiSP 3-D spectral plots of the gradual acceleration from 1-Gz to the 9-Gz level present an interesting panorama of the EEG signal acquired with the dry electrode helmet mounted system. In most subjects, the onset of large-magnitude, low-frequencies in the delta band corresponds to the onset of the subject's straining maneuver to counter the loss of peripheral vision during high Gz-levels.

The onset of each subject's straining maneuver was extracted from the timing and Gz-levels which were simultaneously recorded on the video tape taken during the runs. The video tapes are primarily used to review the subject's response at the various Gz-levels. The 3-D spectral plot of the EEG has the corresponding 3-D spectral plot of the acceleration Gz-level on the y-axis. The spectrum of the accelerometer signal shows the results to be a linearly increasing magnitude (increasing Gz-level) at DC or at the frequency equal to zero Hz. Table II compares the Gz-level at the onset of straining as measured visually verses the Gz-level at the onset of straining as measured from the dry electrode and wet electrode EEG 3-D spectral plots.

TABLE II

COMPARISON OF ONSET OF STRAINING

Gz-level at Onset of Straining During 1-Gz to 9-Gz Acceleration

SUBJECT	VIDEO TAPE	3-D DRY EEG PLOT	3-D WET EEG PLOT
A1	4.86	5.0	N/A
D1	6.73	6.8	6.5
J1	6.7	3.0 **	3.8 **/6.7
L1	5.85	3.7/6.8	5.8
S1	7.18	7.2	3.2/7.8

NOTE: ** subject was talking from 2.8-Gz to 3.9-Gz

In three subjects (Table I: A1, D1, and S1), the Gz-levels as determined from the video recordings and the 3-D dry electrode EEG spectral plots are in agreement with the acceleration level at which the subjects began straining maneuvers, which may infer that significant motion artifacts or low frequency EMG is being seen in the EEG signal. The instantaneous jump in the low frequency activities (less than 4-Hz) that coincides with the beginning of the M-1 straining maneuver can easily be determined from subject's (A-1) 3-D dry electrode EEG spectral plots. The same rapid rise in EEG low frequency activities is not as obvious from the subject's 3-D wet EEG spectral plot to determine when the subject began the M-1 straining maneuver. No G-LOC, nor any antecedent behavioral events could be seen precluding any comments on electrophysiological correlates of G-LOC.

Conclusions And Recommendations of First Study

A fundamental problem with the dry electrode system is the susceptibility to motion artifacts that results in large-magnitude, low frequency signals in the delta and theta bands. Another complication arises from the placement of the EEG electrode leads; placing dry electrodes on the forehead near the temple results in noise or artifacts in the EEG signals that are not brain electrical activities; for example, eye blinks and ocular movements contaminate the lower magnitude EEG signals.

Similarly, placement of the reference electrode (in recordings of EEG with the wet electrodes) over the masseter muscle of the jaw introduces EMG signals into the EEG recordings. The high-frequencies (24 to 36 Hz) observed in the 3-D wet electrode EEG spectral plot of subject S1, which are not as noticeable in the 3-D dry electrode EEG spectral plot of subject S1 may in part be accounted for by the placement of the reference electrode.

Under an optimal environment, e.g., controlled laboratory conditions without motion, the EEG recordings from the dry and wet systems may be similar or well correlated; however, in the centrifuge environment, the spectral plots and bar graphs of the EEG show high frequencies about 24 to 36 Hz, which correspond more with EMG frequencies than with normal EEG frequencies. Studies to answer the question, "Are these signals predominantly, EEG or EMG?" have yet to be conducted.

Regardless of whether the EEGs from the dry electrode system are the same or different as those from the wet electrode system, or whether the signal is really EEG or EMG, a pertinent question is, "Can an algorithm be developed to accurately detect G-LOC?" If the answer is yes, the next question should be, "Can an algorithm be developed to accurately predict the onset of G-LOC in a pilot?"

It is recommended that a study be conducted to evaluate EEG and EMG signals from alternate electrode placements in order to characterize parameters which may serve as indicators of G-induced loss of consciousness. The following suggestions are made for a follow-on study.

1. Record EEG signals from the occipital and/or parietal areas of the head instead of the forehead.

2. Use standard EEG electrodes instead of the ECG recessed paste electrodes

3. Record EMG signals from the abdominal or the back of the leg (gastrocnemius) muscles using standard EMG paste electrodes.

4. Record EEG signals from the frontal area of the head (same locations as the helmet mounted electrodes) using the dry electrode (without a helmet), as well as, wet electrodes for comparison.

5. Do not record the EOG, but use the video recordings to annotate eye-blink times for possible identification and removal from the EEG signals.

SECOND STUDY

Data Acquisition

Volunteers for this study were selected from United States Air Force (USAF) personnel engaged in centrifuge training at Brooks AF base. Since the subjects are in different stages of training and the study is to be conducted on a non-interference with the subject's training; the acceleration profile or protocol for any subject is based on the individual's training requirement. The acceleration-profiles in this study could include any set of rapid onsets to 2-Gz, 4.5-Gz, 5-Gz, 7-Gz, 8-Gz and/or 9-Gz for varying duration of time (subject//training dependent). More advanced trainees experience the Standard Air Combat Maneuver (SACM) which subjects the individual to several consecutive exposures to 4.5-Gz for 15 seconds followed immediately, by a rapid onset to 7-Gz for 15 seconds. The SACM sequence of changing Gz-levels is continued until either the subject, medical monitor, or the centrifuge controller terminates the run.

Electrode Montage

Standard Grass Instruments electroencephalograph (EEG) electrodes were used to record the channels of EEG from the occipital region (O1) to left mastoid (A1), the parietal region (P3) to left mastoid (A1), and the frontal region (between the first point; FP1 and the frontal F7) to left mastoid (A1) with a common reference electrode on the right mastoid (A2). Standard recessed electromyogram (EMG) electrodes were placed on the abdominal muscles, to record EMG during straining maneuvers. Additionally, a single channel of acceleration or Gz-levels of the acceleration profile was recorded.

Recording Equipment

Signals from the EEG and EMG electrodes were input to four differential instrumentation amplifiers (DATA, Inc., Model 2421). The amplifiers were mounted on a plate and positioned inside the centrifuge gondola, aft and to the right of the subject's head. The instrumentation amplifiers which were purchased in 1982 were

23-13

designed to operate from internal nickel-cadmium battery packs. The units can function in either DC or AC mode. In this study the amplifiers were set to AC-coupled, 0.5 Hz high-pass filter cutoff frequency, 300 Hz low-pass filter cutoff frequency, and 10,000 (104) gain. Although the common mode rejection ratio (CMS-defined as the ratio of the differential gain to the common mode gain) was designed to be 80 dB at 60 Hz with either AC or DC coupling, the CMRR was less than 60 dB in all amplifiers.. Signals from the amplifier outputs were transmitted via RG-59 cables, through the centrifuge slip-rings to the patch-panel in the control room. From the patch panel, signals were routed to a RACAL (Model "V-Store" - 16-channel, VHS, magnetic-tape recorder) for storage and subsequent analysis. These amplifiers are old and unstable with tendencies to drift erratically, thus resulting in the loss of frontal EEG and abdominal EMG records

Data Conversion Process

Before spectral analysis could be performed on the recorded EEG signals, the analog signal were converted to digital signals. The process requires a VHS-tape unit, which for our study is a second RACAL recorder. The output signals from the RACAL are input through a bank of programmable low-pass filters (Precision Filters, Inc. Model PS1-1). These low-pass filters, which serve as anti-alias filters, were set to unity gain and cutoff frequencies of 50 Hz. All signal connections were made with standard BNC-connectors and RG-59 cables.

The filtered signals are input to a LabVIEW 16-channel data acquisition board (12-bit analog-to-digital [A/D] converter) via an in-house terminal input board. The analog-to-digital conversion was facilitated by use of LabVIEW's "Daq"-data acquisition software which permits software programming of the, sampling rates, sampling duration, and number of desired channels to digitize. For this study, five data channels were set as:

Channel 0	-	Occipital EEG (O1-A1)
Channel 1	-	Parietal EEG (P3-A1)
Channel 2	-	Frontal EEG (F7-A1)
Channel 3	-	EMG
Channel 4	-	Gz

The input voltage was set to the range, - 1.25 to + 1.25, and the sampling rate was set to 128 samples per second.

Digitized data were stored in a directory called MAC-DATA.

After the data are digitized an in-house file formatting program called "Wmac2ibm" is used to convert the LabVIEW digitized data into a usable format for the DADiSP program (Digital Analysis and display software from DSP Development Corporation). The Wmac2ibm conversion program was written in Pearl Script by Mr. Curtis White, AL/CFT, Brooks AFB, TX. The output of this program, a header file for the DADiSP and the converted digital data files, are stored in a directory called DSP-DATA. In the conversion, the digitized data are concatenated to form a single series file.

Digital Signal Processing

The converted digital data files for DADiSP are then imported into the DADiSP Program as "Series" before the program can process the EEG data. Several windows must be opened before the imported data can be entered into a DADiSP window for processing. This is accomplished by opening a series under "FILE;" "OPEN"; "SERIES" in the pull down menu. The series were opened in the following order:

Window 1, Gz Acceleration,

Window 2, EEG Occipital-O1,

Window 3, EEG Parietal-P,

Window 4, EEG Frontal,

Window 5, EMG.

The advantage of using the DADiSP program is that it permits extraction of data segments from a data file via a cursor macro. This feature is particularly useful with the varying acceleration profiles. For this study, 12-second segments at constant acceleration are extracted from each EEG-data series. The acceleration (Gz) data series is used to determine cursor start and stop timing points which are used in extraction of segments from the three EEG series that correspond to the data between cursor points. The processing of an EEG series includes the following steps:

1. Extraction of a segment which corresponds to a constant acceleration level,

2. Removing any DC-offset from the extracted data,

3. Calibration and scaling,

4. Windowing the data with a Kaiser Window for side-lobe suppression in the spectrum,

5. Calculating the power spectra via a mixed-radix Fast Fourier Transform (FFT) algorithm,

6. Hanning Smoothing to obtain a better spectral estimate,

7. Normalization and extraction of the spectrum from DC to 32 Hz (frequencies of most interest in EEG analysis),

8. Raveling the 32-Hz Bandwidth Spectra into eight, 4-Hz bands which correspond to the standard clinical EEG frequency bands,; i.e., Delta (1-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta 1 (12-16 Hz), etc. to Beta 5 (28-32 Hz).

The "raveled" data is then written as an "ASCII, TABLE-file" to a directory on the computer hard drive.

EEG Parameter-Extraction Process

The Microsoft EXCEL software was used to extract EEG parameters or features from the raveled EEG data. A macro was written to calculate the following 14 features from each file:

(a) Normalized (relative) power in each of the eight 4-Hz EEG clinical bands.

(b) The mean or centroid of the frequency distribution within a band. This parameter is calculated by the

equation: $\qquad MPF = (\sum Ci * Fi) / (\sum Ci) \qquad$ in Hz.

The mean frequency (MPF) was calculated for the three bands of most interest; i.e., Delta, Theta, and Alpha.

(c) The peak or mode of the frequency distribution within a band. The peak frequency (PKF) was obtained by sorting the data within a band for its maximum value and recording the index or frequency at which the maximum value occurred. The PKFs were determined for the three lower EEG frequency bands, Delta, Theta and Alpha.

EXCEL spreadsheets were used to organize the extracted parameters in summary files. Summary data were used to produce 3-D bar charts to show variations of clinical EEG frequency bands with Gz-levels. The most useful set of parameters was found to be the relative power, which corresponds to the normalized power spectra of the EEG.

Results and Discussion of Second Study

Prior to recording from volunteer subjects undergoing acceleration profiles, initial tests were performed to insure system integrity and accuracy. The static test included a system calibration test in which a 10 Hz, 100 microvolt sinusoidal waveform was transmitted through the EEG instrumentation amplifiers, centrifuge slip-rings, patch-panel, and low-pass filters to the Racal magnetic tape recorder.

Additionally, about two minutes of the EEG channels were recorded from a volunteer. This subject was asked to sit and read a book in the centrifuge gondola while the centrifuge was stopped (not in operation) and we recorded the three channels of EEG. The recorded data were analyzed and a DADiSP 3-D spectral plot.

Only two subjects were recorded and analyzed in this preliminary study. The first subject, A, is an experienced centrifuge subject (male), who was riding the centrifuge to remain current. The second subject was an inexperienced subject (female), who had on a previous occasion experience G-LOC. Subject A was subjected to a two-minute SACM profile, but subject B was not.

It was observed in the DADiSP 3-D plot of the EEG activity during the SACM, that subject A's anti-G countering maneuver at the first 7-G encounter was maximal strain. The maximal anti-G countering strain which was very noticeable, resulted in very high levels of physiological electrical activity in the 16 to 32 Hz range. In subsequent +7-Gz accelerations less activity is observed in the 12 to 28 HZ range. The same information is identified in the EXCEL 3-D bar. graph shown in Figure 2. The difference between the 3-D spectral graph and the 3-D bar graph is that the bar graphs represent the summation of normalized power within a 4-Hz spectral band for each 12 seconds of EEG data, whereas, the 3-D spectral graph is the result of moving a spectral window in time. In this case, the spectral window is moved so that there will be a 75% overlap of the previous window.

The results from the inexperienced centrifuge rider (subject B) indicate unusually high activity during the 15 seconds of sustained (constant) Gz-levels of 5 and 7-Gs. Figure 3 shows elevated EEG activities in the 12 to 20 Hz range during 5-Gz acceleration and Beta 1 (12-16 Hz) during the 7-Gz acceleration. An evaluation of the EEG spectra at 7-Gz shows a very high peak at 13.6 Hz on all EEG channels; however, it is most prominent in the parietal area (40% of spectral power) as compared to the occipital area (34% of spectral power).

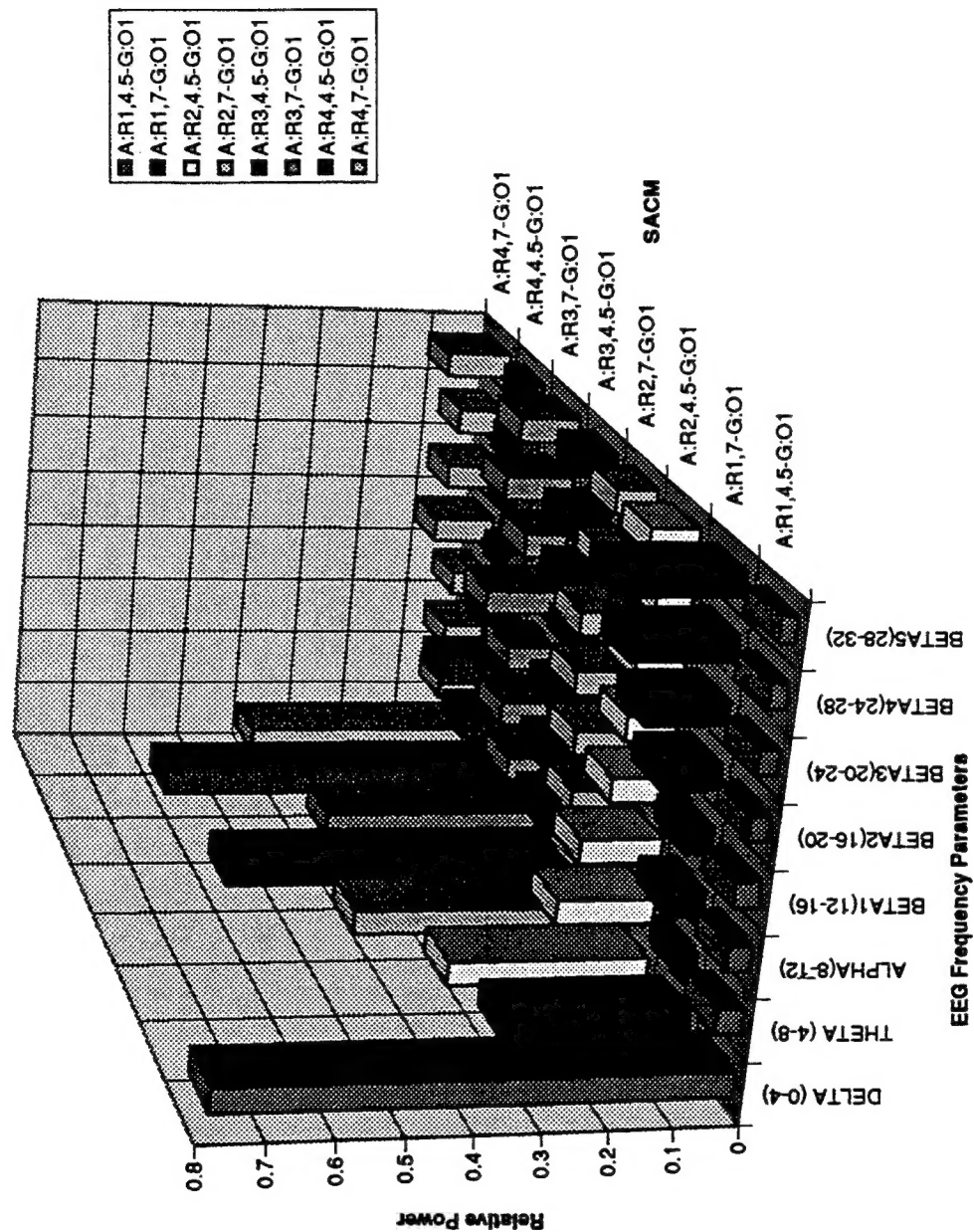The concentration of activities centered about 13.6 Hz is difficult to explain, One can rationalize that it is

Figure 2. 3-D Bar Chart shows the relative power in the clinical EEG bands with increasing +Gz acceleration levels during a Standard Air Combat Maneuvers. Most noticeable is the increase in relative power of frequencies in the 20 - 32 Hz range during the first 15 second at +7 Gz acceleration.
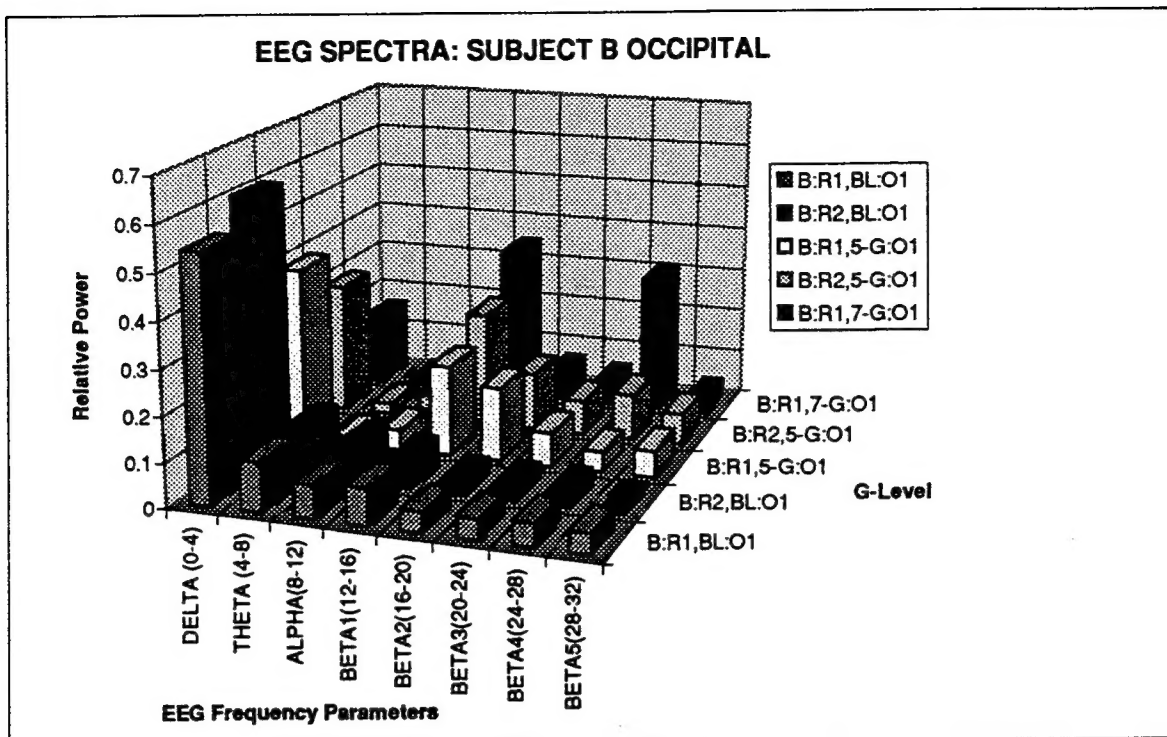
Figure 3 a. 3-D Bar Chart of relative power of clinical EEG bands from the occipital area (O1).
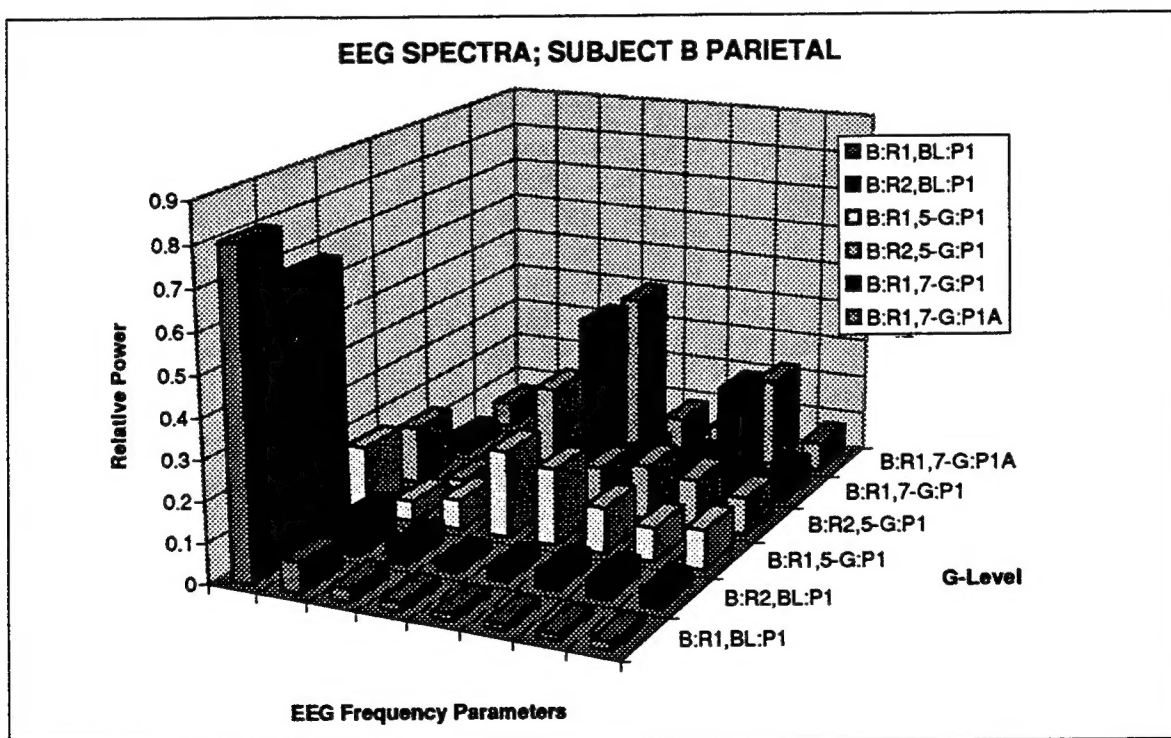

Figure 3 b. 3-D Bar Chart shows relative power in the clinical EEG bands from the parietal area (P3). Most noticable in 3a and 3b is the highly concentrated relative power in the Beta 1 band (12 - 16 Hz) and the Beta 4 band (24 - 28 Hz) at +7 Gz constant acceleration.

an artifact of the centrifuge slip rings; however, one can also speculate that it might be EEG activity since sleep spindles in stage 2 have a frequency about 13.5 Hz. The fact that the signal is stronger in the parietal region than in the occipital or frontal areas is in agreement with the fact that sleep spindles are most prominent in the central ad parietal regions. One can rule out aliasing of the spectrum as the sampling rate of 128-Hz, and the anti-aliasing filtering prior to digitizing preclude this reason as a possibility.

Since the centralization or organization of activity was not observed in the two 5-Gz runs and only one run was made at 7-Gz, it is better to report this as an observation until similar findings can be obtained from additional subjects or repeated +7-Gz runs on the same subject. Nevertheless, in general, the electrical activity of the electroencephalogram increases as the subject counters the acceleration forces with muscular contraction. The electrode placements in the occipital and parietal areas referenced to the mastoids was selected on the basis of minimizing ocular and EMG artifacts while maximizing the acquisition of the EEG signals. The increase in frequency of the bioelectrical activity, between 16 and 32 Hz, is often attributed to EMG-activity; however, EEG activity in the same range of frequencies are reported in the literature.

Recommendation

I recommend that newer and more reliable EEG preamplifiers are obtained commercially to insure the quality of the EEG signals being acquired. EMG signals from the abdominal and gastrocnemius muscles should be recorded. I strongly recommend that the collection of EEG and EMG from subjects in centrifuge training training should continue.

REFERENCES

1. Burns, J.W., Werchan, P.M., Fanton, J.W., and Dollins, A.B., "Performance Recovery in the Baboon Following +Gz-Induced Loss of consciousness," Technical Report: AL/CF-TR-1993-0110, AF/CFT, Brooks, AFB, TX., Dec. 1993.

2. Gillingham, K.K., "High G Stress and Orientational Stress: Physiologic Effects of Aerial Maneuvering", Aviation, Space, and Environmental Medicine, Nov., pp A10-A20, 1988.